Gap Universality of Generalized Wigner and β-Ensembles

László Erdős[1*] Horng-Tzer Yau[2†]

Institute of Mathematics, University of Munich,
Theresienstrasse 39, D-80333 Munich, Germany
lerdos@math.lmu.de [1]

Department of Mathematics, Harvard University
Cambridge MA 02138, USA
htyau@math.harvard.edu [2]

Nov 15, 2012

Abstract

We consider generalized Wigner ensembles and general β-ensembles with analytic potentials for any $\beta \geq 1$. The recent universality results in particular assert that the local averages of consecutive eigenvalue gaps in the bulk of the spectrum are universal in the sense that they coincide with those of the corresponding Gaussian β-ensembles. In this article, we show that local averaging is not necessary for this result, i.e. we prove that the single gap distributions in the bulk are universal.

AMS Subject Classification (2010): 15B52, 82B44

Keywords: β-ensembles, Wigner-Dyson-Gaudin-Mehta universality, gap distribution, log-gas.

Not for Circulation

*Partially supported by SFB-TR 12 Grant of the German Research Council
†Partially supported by NSF grants DMS-0757425, 0804279

Contents

1 Introduction

The fundamental vision that random matrices can be used as basic models for large quantum systems was due to E. Wigner [53]. He conjectured that the eigenvalue gap distributions of large random matrices were universal ("Wigner surmise") in the sense that large quantum systems and random matrices share the same gap distribution functions. The subsequent works of Dyson, Gaudin and Mehta clarified many related issues regarding this assertion and a thorough understanding of the Gaussian ensembles has thus emerged (see the classical book of Mehta [40] for a summary). There are two main categories of random matrices: the invariant and the non-invariant ensembles. The universality conjecture, which is also known as the Wigner-Dyson-Gaudin-Mehta (WDGM) conjecture, asserts that for both ensembles the eigenvalue gap distributions are universal up to symmetry classes. For invariant ensembles, the joint distribution function of the eigenvalues can be expressed explicitly in terms of one dimensional particle systems with logarithmic interactions (i.e., log-gases) at an inverse temperature β. The values $\beta = 1, 2, 4$, correspond to the classical orthogonal, unitary and symplectic ensembles, respectively. Under various conditions on the external potential, the universality for the classical values $\beta = 1, 2, 4$ was proved, via analysis on the corresponding orthogonal polynomials, by Fokas-Its-Kitaev [31], Deift *et. al.* [11, 14, 15], Bleher-Its [6], Pastur-Shcherbina [43, 44] and in many consecutive works, see e.g. [12, 13, 38, 45, 52]. For nonclassical values of β there is no matrix ensemble behind the model, except for the Gaussian cases [17] via tridiagonal matrices. One may still be interested in the local correlation functions of the log-gas as an interacting particle system. The orthogonal polynomial method is not applicable for nonclassical values of β even for the Gaussian case. Nevertheless, the local statistics can still be described very precisely in the Gaussian case with a different method [50, 51]. The universality for general β-ensembles was established only very recently [7, 8] by a new method based on dynamical methods using Dirichlet form estimates from [21, 22]. This method is important for this article and we will discuss it in more details later on. All previous results achieved by this method, however, required in their statement to consider a local average of consecutive gaps. In the current paper we will prove universality of *each single* gap in the bulk.

Turning to the non-invariant ensembles, the most important class is the $N \times N$ Wigner matrices characterized by the independence of their entries. In general, there is no longer an explicit expression for the joint distribution function for the eigenvalues. However, there is a special class of ensembles, the Gaussian divisible ensembles, that interpolate between the general Wigner ensembles and the Gaussian ones. For these ensembles, at least in the special Hermitian case, there is still an explicit formula for the joint distribution of the eigenvalues based upon the Harish-Chandra-Itzykson-Zuber integral. This formula was first put into a mathematically useful form by Johansson [36] (see also the later work of Ben Arous-Péche [5]) to prove the universality of Gaussian divisible ensembles with a Gaussian component of size order one. In [19], the size of the Gaussian component needed for proving the universality was greatly reduced to $N^{-1/2+\varepsilon}$. More importantly, the idea of approximating Wigner ensembles by Gaussian divisible ones was first introduced and, after a perturbation argument, this resulted in the first proof of universality for Hermitian ensembles with general smooth distributions for matrix elements. The smoothness condition was later on removed in [46, 20].

In his seminal paper [18], Dyson observed that the eigenvalue distribution of Gaussian divisible ensembles is the same as the solution of a special system of stochastic differential equations, commonly known now as the Dyson Brownian motion, at a fixed time t. For short times, t is comparable with the variance of the Gaussian component. He also conjectured that the time to "local equilibrium" of the Dyson Brownian motion is of order $1/N$, which is then equivalent to the

universality of Gaussian divisible ensembles with a Gaussian component of order slightly larger than $N^{-1/2}$. Thus the work [19] can be viewed as proving Dyson's conjecture for the Hermitian case. This method, however, completely tied with an explicit formula that is so far restricted to the Hermitian case.

A completely analytic approach to estimate the time to local equilibrium of the Dyson's Brownian motion was initiated in [21] and further developed in [22, 26, 25], see [30] for a detailed account. In these papers, Dyson's conjecture in full generality was proved [26] and universality was established for generalized Wigner ensembles for all symmetric classes. The idea of a dynamical approach in proving universality turns out to be a very powerful one. Dyson Brownian motion can be viewed as the natural gradient flow for Gaussian β log-gases (we will often use the terminology β log-gases for the β-ensembles to emphasize the logarithmic interaction). The gradient flow can be defined with respect to all β log-gases, not just the Gaussian ones. Furthermore, one can consider gradient flows of local log-gases with fixed "good boundary conditions". Here "local" refers to Gibbs measures on $N^a, 0 < a < 1$, consecutive points of a log-gas with the locations of all other points fixed. By "good boundary conditions" we mean that these external points are *rigid*, i.e. their locations are close to their classical locations given by the limiting density of the original log-gas. Using this idea, we have proved the universality of general β-ensembles in [7, 8] for analytic potential.

The main conclusion of these works is that the local gap distributions of either the generalized Wigner ensembles (in all symmetry classes) or the general β-ensembles are universal in the bulk of the spectrum (see [29] for a recent review). The dynamical approach based on Dyson's Brownian motion and related flows also provides a conceptual understanding for the origin of the universality. For technical reasons, however, these proofs apply to averages of consecutive gaps, i.e. cumulative statistics of N^{ε} consecutive gaps were proven to be universal. Averaging the statistics of the consecutive gaps is equivalent to averaging the energy parameter in the correlation functions. Thus, mathematically, the results were also formulated in terms of universality of the correlation functions with averaging in an energy window of size $N^{-1+\varepsilon}$.

The main goal of this paper is to remove the local averaging in the statistics of consecutive gaps in our general approach using Dyson Brownian motion for both invariant and non-invariant ensembles. We will show that the distribution of *each single* gap in the bulk is universal, which we will refer to as the *single gap universality* or simply the *gap universality* whenever there is no confusion. The single gap universality was proved for a special class of Hermitian Wigner matrices with the property that the first four moments of the matrix elements match those of the standard Gaussian random variable [47] and no other results have been known before. In particular, *the single gap universality has not been proved even for the Gaussian orthogonal ensemble (GOE).*

The gap distributions are closely related to the correlation functions which were often used to state the universality of random matrices. These two concepts are equivalent in a certain average sense. However, there is no rigorous relation between correlation functions at a fixed energy and single gap distributions. Thus our results on single gap statistics do not automatically imply the universality of the correlation functions at a fixed energy which was rigorously proved only for Hermitian Wigner matrices [19, 46, 20, 30].

The removal of a local average in the universality results proved in [21, 27, 28] is a technical improvement in itself and its physical meaning is not especially profound. Our motivation for taking seriously this endeavor is due to that the single gap distribution may be closely related to the distribution of a single eigenvalue in the bulk of the spectrum [33] or at the edge [48, 49]. Since our approach does not rely on any explicit formula involving Gaussian matrices, some extension of this method may provide a way to understand the distribution of an individual eigenvalue of

Wigner matrices.

The main new idea in this paper is an analysis of the Dyson Brownian motion via parabolic regularity using the De Giorgi-Nash-Moser idea. Since the Hamiltonians of the local log-gases are convex, the correlation functions can be re-expressed in terms of a time average of certain random walks in random environments. The connection between correlation functions of general log-concave measures and random walks in random environments was already pointed out in the work of Helffer and Sjöstrand [35] and Naddaf and Spencer [41]. This connection was used as an effective way to estimate correlation functions for several models in statistical physics, see, e.g. [2, 16, 32, 34].

In this paper we observe that the single gap universality is a consequence of the Hölder regularity of the solutions to these random walk problems. Due to the logarithmic interaction, the random walks are long ranged and their rates may be singular. Furthermore, the random environments themselves depend on the gap distributions, which were exactly the problems we want to analyze! If we view these random walks as (discrete) parabolic equations with random coefficients, we find that they are of divergence form and are in the form of the equations studied in the fundamental paper by Caffarelli, Chan and Vasseur [10]. The main difficulty to apply [10] to gain partial regularity is that the diffusion coefficients in our settings are random and can be much more singular than were allowed in [10]. For controlling the singularities of these coefficients, we prove an optimal level repulsion estimate for the local log-gases. With these estimates, we are able to extend the method of [10] to prove Hölder regularity for the solution to these random walks problems. This shows that the single gap distributions are universal for local log-gases with good boundary conditions, which is the key result of this paper.

For β-ensembles, it is known that the rigidity of the eigenvalues ensures that boundary conditions are good with high probability. Thus we can apply the local universality of single gap distribution to get the single gap universality of the β-ensembles. We remark, however, that the current result holds only for $\beta \geq 1$ in contrast to $\beta > 0$ in [7, 8], since the current proof heavily relies on the dynamics of the gradient flow of local log-gases. For non-invariant ensembles, a slightly longer argument using the local relaxation flow is needed to connect the local universality result with that for the original Wigner ensemble. This will be explained in Section 5.

In summary, we have recast the question of the single gap universality for random matrices, envisioned by Wigner in the sixties, into a problem concerning the partial regularity of a parabolic equation in divergence form studied by De Giorgi-Nash-Moser. Thanks to the insight of Dyson and the important progress by Caffarelli-Chan-Vasseur [10], we are able to establish the WDGM universality conjecture for each individual gap via De Giorgi-Nash-Moser's idea. We now introduce our models rigorously and state the main results.

2 Main results

We will have two related results, one concerns the generalized Wigner ensembles, the other one the general β-ensembles. We first define the generalized Wigner ensembles. Let $H = (h_{ij})_{i,j=1}^N$ be an $N \times N$ hermitian or symmetric matrix where the matrix elements $h_{ij} = \bar{h}_{ji}$, $i \leq j$, are independent random variables given by a probability measure ν_{ij} with mean zero and variance $\sigma_{ij}^2 \geq 0$;

$$\mathbb{E}\, h_{ij} = 0, \qquad \sigma_{ij}^2 := \mathbb{E}|h_{ij}|^2. \tag{2.1}$$

The distribution ν_{ij} and its variance σ_{ij}^2 may depend on N, but we omit this fact in the notation. We also assume that the normalized matrix elements have a uniform subexponential decay,

$$\mathbb{P}(|h_{ij}| > x\sigma_{ij}) \leq \theta_1 \exp(-x^{\theta_2}), \qquad x > 0, \tag{2.2}$$

with some fixed constants $\theta_1, \theta_2 > 0$, uniformly in N, i, j. In fact, with minor modifications of the proof, an algebraic decay

$$\mathbb{P}(|h_{ij}| > x\sigma_{ij}) \leq C_M x^{-M}$$

with a large enough M is also sufficient.

Definition 2.1 ([24]) *The matrix ensemble H defined above is called generalized Wigner matrix if the following assumptions hold on the variances of the matrix elements* (2.1)

(A) *For any j fixed*

$$\sum_{i=1}^{N} \sigma_{ij}^2 = 1\,. \tag{2.3}$$

(B) *There exist two positive constants, C_{inf} and C_{sup}, independent of N such that*

$$\frac{C_{inf}}{N} \leq \sigma_{ij}^2 \leq \frac{C_{sup}}{N}. \tag{2.4}$$

Let $\mathbb{P}$ and $\mathbb{E}$ denote the probability and the expectation with respect to this ensemble.

We will denote by $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ the eigenvalues of H. In the special case when $\sigma_{ij}^2 = 1/N$ and h_{ij} is Gaussian, the joint probability distribution of the eigenvalues is given

$$\mu = \mu_G^{(N)}(\mathrm{d}\boldsymbol{\lambda}) = \frac{e^{-N\beta\mathcal{H}(\boldsymbol{\lambda})}}{Z_\beta}\mathrm{d}\boldsymbol{\lambda}, \qquad \mathcal{H}(\boldsymbol{\lambda}) = \sum_{i=1}^{N} \frac{\lambda_i^2}{4} - \frac{1}{N}\sum_{i<j} \log|\lambda_j - \lambda_i|. \tag{2.5}$$

The value of β depends on the symmetry class of the matrix; $\beta = 1$ for GOE, $\beta = 2$ for GUE and $\beta = 4$ for GSE. Here Z_β is the normalization factor so that μ is a probability measure.

It is well known that the density or the one point correlation function of μ converges, as $N \to \infty$, to the Wigner semicircle law

$$\varrho(x) := \frac{1}{2\pi}\sqrt{(4 - x^2)_+}. \tag{2.6}$$

We use the notation γ_j for the j-th quantile of this density, i.e. γ_j is defined by

$$\frac{j}{N} = \int_{-2}^{\gamma_j} \varrho_G(x)\mathrm{d}x. \tag{2.7}$$

Our main result on the generalized Wigner matrices asserts that the local gap statistics in the bulk of the spectrum are universal for any general Wigner matrix, in particular they coincide with those of the Gaussian case.

For any integers $A < B$ we introduce the notation $[\![A, B]\!] := \{A, A+1, \ldots, B\}$.

Theorem 2.2 (Gap universality for Wigner matrices) *Let H be a generalized Wigner ensemble with subexponentially decaying matrix elements,* (2.2). *Fix a positive number $\alpha > 0$, an*

integer $n \in \mathbb{N}$ and a smooth, compactly supported function $O : \mathbb{R}^n \to \mathbb{R}$. There exists an $\varepsilon > 0$ and $C > 0$, depending only on α and on O such that

$$\Big| \big[\mathbb{E} - \mathbb{E}^\mu\big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \dots, N(x_j - x_{j+n})\big)\Big| \le CN^{-\varepsilon}, \tag{2.8}$$

for any $j \in [\![\alpha N, (1-\alpha)N]\!]$ and for any sufficiently large $N \ge N_0$, where N_0 depends on all parameters of the model, as well as on n and α.

More generally, for any $k, m \in [\![\alpha N, (1-\alpha)N]\!]$ we have

$$\begin{aligned} \Big| &\mathbb{E} O\big((N\varrho_k)(x_k - x_{k+1}), (N\varrho_k)(x_k - x_{k+2}), \dots, (N\varrho_k)(x_k - x_{k+n})\big) \\ &- \mathbb{E}^\mu O\big((N\varrho_m)(x_m - x_{m+1}), (N\varrho_m)(x_m - x_{m+2}), \dots, (N\varrho_m)(x_m - x_{m+n})\big)\Big| \le CN^{-\varepsilon}, \end{aligned} \tag{2.9}$$

where the local density ϱ_k is defined by $\varrho_k := \varrho(\gamma_k)$.

It is well-known that the gap distribution of Gaussian random matrices for all symmetry classes can be explicitly expressed via a Fredholm determinant provided that a certain local average is taken, see [11, 12, 13]. The result for a single gap, i.e. without local averaging, was only achieved recently in the special case of the Gaussian unitary ensemble (GUE) by Tao [47] (which then easily implies the same results for Hermitian Wigner matrices satisfying the four moment matching condition). It is not clear if a similar argument can be applied to the GOE case.

We now define the β-ensembles with a general external potential. Let $\beta > 0$ be a fixed parameter. Let $V(x)$ be a real analytic potential on $\mathbb{R}$ that grows faster than $(2+\varepsilon)\log|x|$ at infinity and satisfies

$$\inf_{\mathbb{R}} V'' > -\infty. \tag{2.10}$$

Consider the measure

$$\mu = \mu_{\beta,V}^{(N)}(\mathrm{d}\boldsymbol{\lambda}) = \frac{e^{-N\beta\mathcal{H}(\boldsymbol{\lambda})}}{Z_\beta}\mathrm{d}\boldsymbol{\lambda}, \qquad \mathcal{H}(\boldsymbol{\lambda}) = \frac{1}{2}\sum_{i=1}^N V(\lambda_i) - \frac{1}{N}\sum_{i<j}\log|\lambda_j - \lambda_i|. \tag{2.11}$$

Since μ is symmetric in all its variables, we will mostly view it as a measure restricted to the simplex

$$\Xi^{(N)} := \{\boldsymbol{\lambda} \,:\, \lambda_1 < \lambda_2 < \ldots < \lambda_N\} \subset \mathbb{R}^N. \tag{2.12}$$

Note that the Gaussian measure (2.5) is a special case of (2.11) with $V(\lambda) = \lambda^2/2$. In this case we use the notation μ_G for μ.

Let

$$\varrho_1^{(N)}(\lambda) := \mathbb{E}^\mu \frac{1}{N}\sum_{j=1}^N \delta(\lambda - \lambda_j)$$

denote the density, or the one-point function, of μ. It is well known [1, 9] that $\varrho_1^{(N)}$ converges weakly to the equilibrium density $\varrho = \varrho_V$ as $N \to \infty$. The equilibrium density can be characterized as the unique minimizer (in the set of probability measures on $\mathbb{R}$ endowed with the weak topology) of the functional

$$I(\nu) = \int V(t)\mathrm{d}\nu(t) - \iint \log|t-s|\mathrm{d}\nu(t)\mathrm{d}\nu(s). \tag{2.13}$$

In the case, $V(x) = x^2/2$, the minimizer is the Wigner semicircle law $\varrho = \varrho_G$, defined in (2.6), where the subscript G refers to the Gaussian case. In the general case we assume that $\varrho = \varrho_V$ is supported on a single compact interval, $[A, B]$ and $\varrho \in C^2(A, B)$. Moreover, we assume that V is *regular* in the sense that ϱ is strictly positive on (A, B) and vanishes as a square root at the endpoints, see (1.4) of [8]. It is known that these condition are satisfied if, for example, V is strictly convex.

For any $j \le N$ define the classical location of the j-th particle $\gamma_{j,V}$ by

$$\frac{j}{N} = \int_A^{\gamma_{j,V}} \varrho_V(x)\mathrm{d}x, \tag{2.14}$$

and for the Gaussian case we have $[A, B] = [-2, 2]$ and we use the notation $\gamma_{j,G} = \gamma_j$ for the corresponding classical location, defined in (2.7). We set

$$\varrho_j^V := \varrho_V(\gamma_{j,V}), \qquad \text{and} \qquad \varrho_j^G := \varrho_G(\gamma_{j,G}) \tag{2.15}$$

to be the limiting density at the classical location of the j-th particle. Our main theorem on the β-ensembles is the following.

Theorem 2.3 (Gap universality for β-ensembles) *Let $\beta \ge 1$ and V be a real analytic potential with* (2.10) *such that ϱ_V is supported on a single compact interval, $[A, B]$, $\varrho_V \in C^2(A, B)$, and that V is regular. Fix a positive number $\alpha > 0$, an integer $n \in \mathbb{N}$ and a smooth, compactly supported function $O : \mathbb{R}^n \to \mathbb{R}$. Let $\mu = \mu_V = \mu_{\beta,V}^{(N)}$ be given by* (2.11) *and let μ_G denote the same measure for the Gaussian case. Then there exist an $\varepsilon > 0$, depending only on α, β and the potential V, and a constant C depending on O such that*

$$\Big|\mathbb{E}^{\mu_V} O\Big((N\varrho_k^V)(x_k - x_{k+1}), (N\varrho_k^V)(x_k - x_{k+2}), \ldots, (N\varrho_k^V)(x_k - x_{k+n})\Big) \tag{2.16}$$

$$- \mathbb{E}^{\mu_G} O\Big((N\varrho_m^G)(x_m - x_{m+1}), (N\varrho_m^G)(x_m - x_{m+2}), \ldots, (N\varrho_m^G)(x_m - x_{m+n})\Big)\Big| \le CN^{-\varepsilon}$$

for any $k, m \in [\![\alpha N, (1-\alpha)N]\!]$ and for any sufficiently large $N \ge N_0$, where N_0 depends on V, β, as well as on n and α. In particular, the distribution of the rescaled gaps w.r.t. μ_V does not depend on the index k in the bulk.

Theorem 2.3, in particular, asserts that the single gap distribution in the bulk is independent of the index k. The special GUE case of this assertion is the content of [47] where the proof uses some special structures of GUE.

The proofs of both Theorems 2.2 and 2.3 rely on the uniqueness of the gap distribution for a localized version of the equilibrium measure (2.5) with a certain class of boundary conditions. This main technical result will be formulated in Theorem 3.1 in the next section after we introduce the necessary notations. An orientation of the content of the paper will be given at the end of Section 3.1.

We remark that Theorem 2.3 is stated only for $\beta \ge 1$; on the contrary, the universality with local averaging in [7, 8] was proved for $\beta > 0$. The main reason is that the current proof relies heavily on the dynamics of the gradient flow of local log-gases. Hence the well-posedness of the dynamics is crucial which is available only for $\beta \ge 1$. On the other hand, in [7, 8] we use only certain Dirichlet form inequalities (see, e.g. Lemma 5.9 in [7]), which we could prove with an effective regularization scheme for all $\beta > 0$. For $\beta < 1$ it is not clear if such a regularization can also be applied to the new inequalities we will prove here.

3 Local equilibrium measures

3.1 Basic properties of local equilibrium measures

Fix two small positive numbers, $\alpha, \delta > 0$. Choose two positive integer parameters L, K such that

$$L \in [\![\alpha N, (1-\alpha)N]\!], \qquad N^\delta \le K \le N^{1/4}. \tag{3.1}$$

We consider the parameters L and K fixed and often we will not indicate them in the notation. All results will hold for any sufficiently small α, δ and for any sufficiently large $N \ge N_0$, where the threshold N_0 depends on α, δ and maybe on other parameters of the model. Throughout the paper we will use C and c to denote positive constants which, among others, may depend on α, δ and on the constants in (2.2) and (2.4), but we will not emphasize this dependence. Typically C denotes a large generic constant, while c denotes a small one whose values may change from line to line. These constants are independent of K and N, which are the limiting large parameters of the problem, but they may depend on each other. In most cases this interdependence is harmless since it only requires that a fresh constant C be sufficiently large or c be sufficiently small, depending on the size of the previously established generic constants. In some cases, however, the constants are related in a more subtle manner. In this case we will use $C_0, C_1, \ldots$ and $c_0, c_1, \ldots$ etc. to denote specific constants in order to be able to refer to them along the proof.

For convenience, we set

$$\mathcal{K} := 2K + 1.$$

Denote $I = I_{L,K} := [\![L-K, L+K]\!]$ the set of $\mathcal{K}$ consecutive indices in the bulk. We will distinguish the inside and outside particles by renaming them as

$$(\lambda_1, \lambda_2, \ldots, \lambda_N) := (y_1, \ldots y_{L-K-1}, x_{L-K}, \ldots, x_{L+K}, y_{L+K+1}, \ldots y_N) \in \Xi^{(N)}. \tag{3.2}$$

Note that the particles keep their original indices. The notation $\Xi^{(N)}$ refers to the simplex (2.12). In short we will write

$$\mathbf{x} = (x_{L-K}, \ldots x_{L+K}), \qquad \text{and} \qquad \mathbf{y} = (y_1, \ldots y_{L-K-1}, y_{L+K+1}, \ldots y_N).$$

These points are always listed in increasing order, i.e. $\mathbf{x} \in \Xi^{(\mathcal{K})}$ and $\mathbf{y} \in \Xi^{(N-\mathcal{K})}$. We will refer to the y's as the *external* points and to the x's as *internal* points.

We will fix the external points (often called boundary conditions) and study the conditional measures on the internal points. We first define the *local equilibrium measure* (or *local measure* in short) on $\mathbf{x}$ with boundary condition $\mathbf{y}$ by

$$\mu_{\mathbf{y}}(\mathrm{d}\mathbf{x}) := \mu_{\mathbf{y}}(\mathbf{x})\mathrm{d}\mathbf{x}, \qquad \mu_{\mathbf{y}}(\mathbf{x}) := \mu(\mathbf{y}, \mathbf{x}) \left[\int \mu(\mathbf{y}, \mathbf{x}) \mathrm{d}\mathbf{x} \right]^{-1}, \tag{3.3}$$

where $\mu = \mu(\mathbf{y}, \mathbf{x})$ is the (global) equilibrium measure (2.11) (we do not distinguish between the measure μ and its density function $\mu(\mathbf{y}, \mathbf{x})$ in the notation). Note that for any fixed $\mathbf{y} \in \Xi^{(N-\mathcal{K})}$, all x_j lie in the *open configuration interval*, denoted by

$$J = J_{\mathbf{y}} := (y_{L-K-1}, y_{L+K+1}).$$

Define

$$\bar{y} := \frac{1}{2}(y_{L-K-1} + y_{L+K+1})$$

to be the midpoint of the configuration interval. We also introduce

$$\alpha_j := \bar{y} + \frac{j-L}{\mathcal{K}+1}|J|, \qquad j \in I_{L,K}, \tag{3.4}$$

to denote the $\mathcal{K}$ equidistant points within the interval J.

For any fixed $L, K, \mathbf{y}$, the equilibrium measure can also be written as a Gibbs measure,

$$\mu_{\mathbf{y}} = \mu^{(N)}_{\mathbf{y},\beta,V} = Z_{\mathbf{y}}^{-1} e^{-N\beta\mathcal{H}_{\mathbf{y}}}, \tag{3.5}$$

with Hamiltonian

$$\begin{aligned} \mathcal{H}_{\mathbf{y}}(\mathbf{x}) :=& \sum_{i\in I} \frac{1}{2} V_{\mathbf{y}}(x_i) - \frac{1}{N} \sum_{\substack{i,j\in I\\ i<j}} \log|x_j - x_i|, \\ V_{\mathbf{y}}(x) :=& V(x) - \frac{2}{N} \sum_{j\notin I} \log|x - y_j|. \end{aligned} \tag{3.6}$$

Here $V_{\mathbf{y}}(x)$ can be viewed as the external potential of a β-log-gas of the points $\{x_i \,:\, i \in I\}$ in the configuration interval J.

Our main technical result, Theorem 3.1 below, asserts that, for K, L chosen according to (3.1), the local gap statistics is essentially independent of V and $\mathbf{y}$ as long as the boundary conditions $\mathbf{y}$ are regular. This property is expressed by defining the following set of "good" boundary conditions with some given positive parameters ν, α:

$$\begin{aligned} \mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\nu,\alpha) :=& \{\mathbf{y} : |y_k - \gamma_k| \le N^{-1+\nu}, \quad k \in [\![\alpha N, (1-\alpha)N]\!] \setminus I_{L,K}\} \\ & \cap \{\mathbf{y} : |y_k - \gamma_k| \le N^{-4/15+\nu}, \quad k \in [\![N^{3/5+\nu}, N - N^{3/5+\nu}]\!]\} \\ & \cap \{\mathbf{y} : |y_k - \gamma_k| \le 1, \ \ k \in [\![1, N]\!] \setminus I_{L,K}\}. \end{aligned} \tag{3.7}$$

In Section 4 we will see that this definition is taylored to the previously proven rigidity bounds for the β-ensemble, see (4.4). The rigidity bounds for the generalized Wigner matrices are stronger, see (5.1), so this definition will suit the needs of both proofs.

Theorem 3.1 (Gap universality for local measures) *Fix $L, \widetilde{L}$ and $\mathcal{K} = 2K+1$ satisfying (3.1) with an exponent $\delta > 0$. Consider two boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}}$ such that the configuration intervals coincide,*

$$J = (y_{L-K-1}, y_{L+K+1}) = (\widetilde{y}_{\widetilde{L}-K-1}, \widetilde{y}_{\widetilde{L}+K+1}). \tag{3.8}$$

We consider the measures $\mu = \mu_{\mathbf{y},\beta,V}$ and $\widetilde{\mu} = \mu_{\widetilde{\mathbf{y}},\beta,\widetilde{V}}$ defined as in (3.5), with possibly two different external potentials V and $\widetilde{V}$. Let $\xi > 0$ be a small constant. Assume that $|J|$ satisfies

$$|J| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\Big(\frac{K^\xi}{N}\Big). \tag{3.9}$$

Suppose that $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$ and that

$$\max_{j\in I_{L,K}} \Big|\mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j\Big| + \max_{j\in I_{\widetilde{L},K}} \Big|\mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} x_j - \alpha_j\Big| \le CN^{-1}K^\xi \tag{3.10}$$

holds. Let the integer number p satisfy $|p| \le K - K^{1-\xi^*}$ for some small $\xi^* > 0$. Then there exists $\xi_0 > 0$, depending on δ, such that if $\xi, \xi^* \le \xi_0$ then for any n fixed and any bounded smooth observable $O : \mathbb{R}^n \to \mathbb{R}$ with compact support we have

$$\Big| \mathbb{E}^{\mu_{\mathbf{y}}} O\big(N(x_{L+p} - x_{L+p+1}), \dots N(x_{L+p} - x_{L+p+n})\big) \tag{3.11}$$

$$- \mathbb{E}^{\widetilde{\mu}_{\tilde{\mathbf{y}}}} O\big(N(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+1}), \dots N(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+n})\big) \Big| \le CK^{-\varepsilon}$$

for some $\varepsilon > 0$ depending on δ, α and for some C depending on O. This holds for any $N \ge N_0$ sufficiently large, where N_0 depends on the parameters ξ, ξ^*, α, and C in (3.10).

In the following two theorems we establish rigidity and level repulsion estimates for the local log-gas $\mu_{\mathbf{y}}$ with good boundary conditions $\mathbf{y}$. While both rigidity and level repulsion are basic questions for log gases, our main motivation to prove these theorems is to use them in the proof of Theorem 3.1. The current form of the level repulsion estimate is new which has not been proved in literature (a much weaker form was proved in (4.11) of [7]). The rigidity estimate was proved for the global equilibrium measure μ in [7]. From this estimate, one can conclude that $\mu_{\mathbf{y}}$ has a good rigidity bound for a set of boundary conditions with high probability w.r.t. the global measure μ. However, we will need a rigidity estimate for $\mu_{\mathbf{y}}$ for a set of $\mathbf{y}$'s with high probability with respect to some different measure, which may be asymptotically singular to μ for large N. For example, in the proof for the gap universality of Wigner matrices such a measure is given by the time evolved measure $f_t\mu$, see Section 5. The following result asserts that a rigidity estimate holds for $\mu_{\mathbf{y}}$ provided that $\mathbf{y}$ itself satisfies a rigidity bound and an extra condition, (3.12), holds. This provides explicit criteria to describe the set of "good" $\mathbf{y}$'s whose measure w.r.t. $f_t\mu$ can then be estimated with different methods.

Theorem 3.2 (Rigidity estimate for local measures) *Let L and K satisfy* (3.1) *with δ the exponent appearing in* (3.1). *Let ξ, α be any fixed positive constants. For $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$ consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in* (3.5) *and assume that*

$$\Big| \mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j \Big| \le CN^{-1}K^{\xi}, \quad j \in I = I_{L,K}, \tag{3.12}$$

is satisfied. Then there are positive constants C, c, depending on ξ, such that for any $k \in I$ and $u > 0$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}\Big(\big| x_k - \alpha_k \big| \ge uK^{\xi}N^{-1} \Big) \le Ce^{-cu^2}. \tag{3.13}$$

Now we state the level repulsion estimates which will be proven in Section 6.2.

Theorem 3.3 (Level repulsion estimate for local measures) *Let L and K satisfy* (3.1) *and let ξ, α be any fixed positive constants. Then for $\mathbf{y} \in \mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$ we have the following estimates:*
i) [Weak form of level repulsion] For any $s > 0$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+1} - x_i \le s/N] \le C\,(Ns)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!] \tag{3.14}$$

and

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+2} - x_i \le s/N] \le C(Ns)^{2\beta+1} \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{3.15}$$

(Here we used the convention that $x_{L-K-1} := y_{L-K-1}, x_{L+K+1} := y_{L+K+1}$.)
ii) [Strong form of level repulsion] Suppose that there exist positive constants C, c such that the following rigidity estimate holds for any $k \in I$:

$$\mathbb{P}^{\mu_{\mathbf{y}}}\Big(|x_k - \alpha_k| \ge C K^{\xi^2} N^{-1}\Big) \le C \exp\left(-K^c\right). \tag{3.16}$$

Then there exists small a constant θ, depending on C, c in (3.16), such that for any $s \ge \exp(-K^\theta)$. we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+1} - x_i \le s/N] \le C\left(K^\xi s \log N\right)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!] \tag{3.17}$$

and

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+2} - x_i \le s/N] \le C\left(K^\xi s \log N\right)^{2\beta+1}, \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{3.18}$$

We remark that the estimates (3.18) and (3.15) on the second gap are not needed for the main proof, we listed them only for possible further reference. The exponents are not optimal; one would expect them to be $3\beta + 3$. With some extra work, it should not be difficult to get the optimal exponents. Moreover, our results can be extended to $x_{i+k} - x_i$ for any k fin We also mention that the assumption (3.16) required in part ii) is weaker than what we prove in (3.13). In fact, the weaker form (3.16) of the rigidity would be enough throughout the paper except at one place, at the end of the proof of Lemma 6.8.

Theorem 3.1 is our key result. In Sections 4 and 5 we will show how to use Theorem 3.1 to prove the main Theorems 2.2 and 2.3. Although the basic structure of the proof of Theorem 2.3 is similar to the one given in [7] where a locally averaged version of this theorem was proved under a locally averaged version of Theorem 3.1, here we have to verify the assumption (3.10) which will be done in Lemma 4.1. The proof of Theorem 2.2, on the other hand, is very different from the recent proof of universality in [27, 28]. This will be explained in Section 5.

The proofs of the auxiliary Theorems 3.2 and 3.3 will be given in Section 6. The proof of Theorem 3.1 will start from Section 6.3 and will continue until the end of the paper. At the beginning of Section 6.3 we will explain the main ideas of the proof. For readers interested in the proof of Theorem 3.1, Sections 4 and 5 can be skipped.

3.2 Extensions and further results

We formulated Theorems 3.1, 3.2 and 3.3 with assumptions requiring that the boundary conditions $\mathbf{y}$ are "good". In fact, all these results hold in a more general setting.

Definition 3.4 *An external potential U of a β-log-gas of K points in a configuration interval $J = (a, b)$ is called K^ξ-regular, if the following bounds hold:*

$$|J| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\Big(\frac{K^\xi}{N}\Big), \tag{3.19}$$

$$U'(x) = \varrho(\bar{y}) \log \frac{d_+(x)}{d_-(x)} + O\Big(\frac{K^\xi}{N d(x)}\Big), \qquad x \in J, \tag{3.20}$$

$$U''(x) \ge \inf V'' + \frac{c}{d(x)}, \qquad x \in J, \tag{3.21}$$

with some positive $c > 0$ and for some small $\xi > 0$, where

$$d(x) := \min\{|x-a|, |x-b|\}$$

is the distance to the boundary of J and

$$d_-(x) := d(x) + \varrho(\bar{y})N^{-1}K^{\xi}, \qquad d_+(x) := \max\{|x-a|, |x-b|\} + \varrho(\bar{y})N^{-1}K^{\xi}.$$

The following lemma, proven in Appendix A, asserts that "good" boundary conditions $\mathbf{y}$ give rise to regular external potential $V_{\mathbf{y}}$.

Lemma 3.5 *Let L and K satisfy* (3.1) *and δ is the exponent appearing in* (3.1). *Then for any $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha/2)$ the external potential $V_{\mathbf{y}}$* (3.6) *on the configuration interval $J_{\mathbf{y}}$ is K^{ξ}-regular;*

$$|J_{\mathbf{y}}| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\Big(\frac{K^{\xi}}{N}\Big), \tag{3.22}$$

$$V'_{\mathbf{y}}(x) = \varrho(\bar{y}) \log \frac{d_+(x)}{d_-(x)} + O\Big(\frac{K^{\xi}}{Nd(x)}\Big), \qquad x \in J_{\mathbf{y}}, \tag{3.23}$$

$$V''_{\mathbf{y}}(x) \geq \inf V'' + \frac{c}{d(x)}, \qquad x \in J_{\mathbf{y}}. \tag{3.24}$$

Inspecting the proofs of Theorems 3.1, 3.2, 6.3 and Theorem 3.3, they use only the property that $V_{\mathbf{y}}$ on $J_{\mathbf{y}}$ is regular. Thus the same proofs are valid for any local β-log-gas with an external potential U, where the conditions $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$ or $\mathbf{y} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$ in these statements are replaced by requiring that U is K^{ξ}-regular or K^{ξ^2}-regular, respectively.

4 Gap universality for β-ensembles: proof of Theorem 2.3

4.1 Rigidity bounds and its consequences

The aim of this section is to use Theorem 3.1 to prove Theorem 2.3. In order to verify the assumptions of Theorem 3.1, we first recall the rigidity estimate w.r.t. μ defined in (2.11). Recall that $\gamma_k = \gamma_{k,V}$ denotes the classical location of the k-th point (2.14). For the case of convex potential, in Theorem 3.1 of [7] it was proved that for any fixed $\alpha > 0$ and $\nu > 0$, there are constants $C_0, c_1, c_2 > 0$ such that for any $N \geq 1$ and $k \in [\![\alpha N, (1-\alpha)N]\!]$,

$$\mathbb{P}^{\mu}\left(|\lambda_k - \gamma_k| > N^{-1+\nu}\right) \leq C_0 \exp\left(-c_1 N^{c_2}\right). \tag{4.1}$$

The same estimate holds also for the non-convex case, see Theorem 1.1 of [8].

Near the spectral edges, a somewhat weaker control was proven, see Lemma 3.6 of [7]: for any $\nu > 0$ there are $C_0, c_1, c_2 > 0$ such that

$$\mathbb{P}^{\mu}\left(|\lambda_k - \gamma_k| > N^{-4/15+\nu}\right) \leq C_0 \exp\left(-c_1 N^{c_2}\right) \tag{4.2}$$

for any $N^{3/5+\nu} \leq k \leq N - N^{3/5+\nu}$, if $N \geq N_0(\nu)$ is sufficiently large. We can choose C_0, c_1, c_2 to be the same in (4.1) and (4.2).

We claim that the estimate (4.2) holds also for the non-convex case. Inspecting the proof of Lemma 3.6 of [7], the convexity of V was used only in the logarithmic Sobolev inequality (LSI)

three times. The first application of LSI was in the analysis of the loop equation. As it was already observed in [8], Lemma 2.1 of [8] can replace LSI for this purpose, in particular Lemma 2.2 of [8] holds without convexity. The second application of LSI was to obtain a concentration estimate on $\lambda_j - \mathbb{E}^\mu \lambda_j$ on scale $N^{-1/2}$ uniformly in j (see (3.12) of [7]). In Lemma 3.6 of [8], however, the equivalence of the original measure μ and its convexified version ν (see Definition 3.4 of [8]) was established. In particular, the same concentration estimate was established for the non-convex measure as well (see (3.22) of [8]) and the analogue of Corollary 3.5 of [7] was also obtained (see (3.24) of [8]). Finally, LSI was used also to control the variance term (3.32) of [8]. Thanks to Lemma 3.6 again, this control can be directly transferred from the convexified measure ν to the original measure μ. Taking these ingredients from [8] into account, the proof of Lemma 3.6 of [7] goes through for the non-convex case as well.

Finally, we have a very weak control that holds for all points (see (1.7) in [8]): for any $C > 0$ there are positive constants C_0, c_1 and c_2 such that

$$\mathbb{P}^\mu \left(|\lambda_k - \gamma_k| > C \right) \le C_0 \exp\left(-c_1 N^{c_2}\right). \tag{4.3}$$

Given C, we can choose the positive constants C_0, c_1, c_2 to be the same in (4.1), (4.2) and (4.3).

The set $\mathcal{R}_{L,K}$ in (3.7) was exactly defined as the set of events that these three rigidity estimates hold. From (4.1), (4.2) and (4.3) we have

$$\mathbb{P}^\mu(\mathcal{R}_{L,K}(\nu, \alpha)) \ge 1 - C_0 \exp\left(-c_1 N^{c_2}\right) \tag{4.4}$$

for any $\nu > 0$, $\alpha > 0$ with some positive constants C_0, c_1, c_2 that depend on ν and α.

Lemma 4.1 *Let L and K satisfy* (3.1) *and δ is the exponent appearing in* (3.1). *Then for any small ξ and α there exists a set $\mathcal{R}^* = \mathcal{R}^*_{L,K,\mu}(\xi^2\delta/2, \alpha/2) \subset \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$ such that*

$$\mathbb{P}^\mu(\mathcal{R}^*) \ge 1 - C_0 \exp\Big(-\frac{1}{2} c_1 N^{c_2}\Big) \tag{4.5}$$

with the constants C_0, c_1, c_2 from (4.1). *Moreover, for any $\mathbf{y} \in \mathcal{R}^*$ we have*

$$\big| \mathbb{E}^{\mu_\mathbf{y}} x_k - \alpha_k \big| \le C N^{-1} K^\xi, \qquad k \in I_{L,K}, \tag{4.6}$$

where α_k was defined in (3.4).

Proof. For any $\nu > 0$ define

$$\mathcal{R}^*_{L,K,\mu}(\nu,\alpha) := \Big\{ \mathbf{y} \in \mathcal{R}(\nu,\alpha) \ : \ \mathbb{P}^{\mu_\mathbf{y}} \left(|x_k - \gamma_k| > N^{-1+\nu}\right) \le \exp\big(-\frac{1}{2} c_1 N^{c_2}\big), \ \ \forall k \in I_{L,K} \Big\} \tag{4.7}$$

with the ν-dependent constants $c_1, c_2 > 0$ from (4.4). Note that $\mathcal{R}^*$, unlike $\mathcal{R}$, depends on the underlying measure μ through the family of its conditional measures $\mu_\mathbf{y}$. Applying (4.4) for $\nu = \xi^2\delta/2$ and setting $\mathcal{R} = \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$, $\mathcal{R}^* = \mathcal{R}^*_{L,K,\mu}(\xi^2\delta/2, \alpha/2)$, we have

$$\mathbb{P}^\mu(\mathcal{R}^*) \ge 1 - C_0 \exp\big(-\frac{1}{2} c_1 N^{c_2}\big)$$

with some C_0, c_1, c_2. Now if $\mathbf{y} \in \mathcal{R}^*$, then

$$\big| \mathbb{E}^{\mu_\mathbf{y}} x_k - \gamma_k \big| \le C_0 e^{-c_1 N^{c_2}/3} + C N^{-1} K^{\xi^2}, \qquad k \in I_{L,K}. \tag{4.8}$$

In order to prove (4.6), it remains to show that $|\alpha_k - \gamma_k|$ is bounded by $CN^{-1}K^{\xi}$ for any $k \in I_{L,K}$. To see this, we can use that $\varrho \in C^1$ away from the edge, thus

$$\varrho(x) = \varrho(\bar{y}) + O(x - \bar{y})$$

(recall that $\bar{y}$ is the midpoint of J). By Taylor expansion we have

$$k-(L-K-1) = N\int_{\gamma_{L-K-1}}^{\gamma_k} \varrho = N\int_{y_{L-K-1}}^{\gamma_k} \varrho + O(N^{\xi\delta/2}) = N|\gamma_k - y_{L-K-1}|\varrho(\bar{y}) + O(N|J|^2 + N^{\xi\delta/2}),$$

i.e.

$$\gamma_k = y_{L-K-1} + \frac{k-L+K+1}{N\varrho(\bar{y})} + O(N^{-1}K^{\xi}). \tag{4.9}$$

Here we used that $J = J_{\mathbf{y}}$ satisfies (3.22) according to Lemma 3.5, since $\mathbf{y} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2) \subset \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Comparing (4.9) with the definition of α_k, (3.4), using (3.22) and the fact that $\bar{y} - y_{L-K-1} = \frac{1}{2}|J|$, we have

$$|\alpha_k - \gamma_k| \le CN^{-1}K^{\xi}. \tag{4.10}$$

Together with (4.8) this implies (4.6) and this completes the proof of Lemma 4.1. □

4.2 Completing the proof of Theorem 2.3

We first notice that it is sufficient to prove Theorem 2.3 for the special case $m = N/2$, i.e. when the local statistics for the Gaussian measure is considered at the central point of the spectrum. Indeed, once Theorem 2.3 is proved for any V, k and $m = N/2$, then with the choice $V(x) = x^2/2$ we can use it to establish that the local statistics for the Gaussian measure around any fixed index k in the bulk coincide with the local statistics in the middle. So from now on we assume $m = N/2$, but we carry the notation m for simplicity.

Given k and $m = N/2$ as in (2.16), we first choose $L, \widetilde{L}, K$, satisfying (3.1) (maybe with a smaller α than given in Theorem 2.3), so that $k = L + p$, $m = \widetilde{L} + p$ hold for some $|p| \le K/2$. In particular

$$|\widetilde{L} - N/2| \le K. \tag{4.11}$$

For brevity, we use $\mu = \mu_V$ and $\widetilde{\mu} = \mu^G$ in accordance with the notation of Theorem 3.1.

We consider $\mathbf{y} \in R^*_{L,K,\mu}(\xi^2\delta/2, \alpha)$ and $\widetilde{\mathbf{y}} \in R^*_{\widetilde{L},K,\widetilde{\mu}}(\xi^2\delta/2, \alpha)$, where δ is the exponent appearing in (3.1). We omit the arguments and recall that

$$\mu(R^*_{L,K,\mu}) \ge 1 - C_0 \exp\Big(-\frac{1}{2}c_1 N^{c_2}\Big), \qquad \widetilde{\mu}(R^*_{\widetilde{L},K,\widetilde{\mu}}) \ge 1 - C_0 \exp\Big(-\frac{1}{2}c_1 N^{c_2}\Big) \tag{4.12}$$

with some positive constants.

Proposition 4.2 *With the above choice of the parameters and for any $\mathbf{y} \in R^*_{L,K,\mu}(\xi^2\delta/2, \alpha)$ and $\widetilde{\mathbf{y}} \in R^*_{\widetilde{L},K,\widetilde{\mu}}(\xi^2\delta/2, \alpha)$, we have*

$$\begin{aligned} \Big| &\mathbb{E}^{\mu_{\mathbf{y}}} O\big((N\varrho^V_{L+p})(x_{L+p} - x_{L+p+1}), \ldots (N\varrho^V_{L+p})(x_{L+p} - x_{L+p+n})\big) \qquad (4.13) \\ &- \mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} O\big((N\varrho^G_{\widetilde{L}+p})(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+1}), \ldots (N\varrho^G_{\widetilde{L}+p})(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+n})\big)\Big| \le CK^{-\varepsilon}, \end{aligned}$$

where ε is from Theorem 3.1.

Theorem 2.3 follows immediately from (4.12) and this proposition. □

The rest of this section is devoted to the proof of Proposition 4.2.

Proof of Proposition 4.2. We will apply Theorem 3.1, but first we have to bring the two measures onto the same configuration interval J to satisfy (3.8). This will be done in three steps. First, using the scale invariance of the Gaussian log-gas we rescale it so that the local density approximately matches with that of μ_V. This will guarantee that the two configuration intervals have almost the same length. In the second step we adjust the local Gaussian log-gas $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ so that $J_{\widetilde{\mathbf{y}}}$ has exactly the correct length. Finally, we shift the two intervals so that they coincide. This allows us to apply Theorem 3.1 to conclude the local statistics are identical.

The local densities ϱ_V around $\gamma_{L+p,V}$ and ϱ_G around $\gamma_{\widetilde{L}+p,G}$ may considerably differ. So in the first step we rescale the Gaussian log-gas so that

$$\varrho_V(\gamma_{L+p,V}) = \varrho_G(\gamma_{\widetilde{L}+p,G}). \tag{4.14}$$

To do that, recall that we defined the Gaussian log-gas with the standard $V(x) = x^2/2$ external potential, but we could choose $V_s(x) = s^2x^2/2$ with any fixed $s > 0$ and consider the Gaussian log-gas

$$\mu_G^s(\boldsymbol{\lambda}) \sim \exp\big(-N\beta\mathcal{H}_s(\boldsymbol{\lambda})\big), \qquad \mathcal{H}_s(\boldsymbol{\lambda}) := \frac{1}{2}\sum_{i=1}^N V_s(\lambda_i) - \frac{1}{N}\sum_{i<j}\log|\lambda_j - \lambda_i|.$$

This results in a rescaling of the semicircle density ϱ_G to $\varrho_G^s(x) := s\varrho_G(sx)$ and $\gamma_{i,G}$ to $\gamma_{i,G}^s := s^{-1}\gamma_{i,G}$ for any i, so $\varrho_G(\gamma_{i,G})$ gets rescaled to $\varrho_G^s(\gamma_{i,G}^s) = s\varrho_G(\gamma_{i,G})$. In particular, $\varrho_G(\gamma_{\widetilde{L}+p,G})$ is rescaled to $s\varrho_G(\gamma_{\widetilde{L}+p,G})$, and thus choosing s appropriately, we can achieve that (4.14) holds (keeping the left hand side fixed). Set

$$\mathcal{O}_s(\mathbf{x}) := O\big((N\varrho_G^s(\gamma_{m,G}^s))(x_m - x_{m+1}), \ldots, (N\varrho_G^s(\gamma_{m,G}^s))(x_m - x_{m+n})\big), \qquad m = \widetilde{L} + p,$$

and notice that $\mathcal{O}_s(\mathbf{x}) = \mathcal{O}(s\mathbf{x})$. This means that the local gap statistics $\mathbb{E}^{\mu_G^s}\mathcal{O}_s$ is independent of the scaling parameter s, since the product $(N\varrho_m^G)(x_m - x_{m+a})$ (notation defined in (2.15)) is unchanged under the scaling. So we can work with the rescaled Gaussian measure. For notational simplicity we will not carry the s parameter further and we just assume that (4.14) holds with the original Gaussian $V(x) = x^2/2$.

We have now achieved that the two densities at at some points of the configuration intervals coincide, but the lengths of these two intervals still slightly differ. In the second step we match them exactly. Since $\mathbf{y} \in R_{L,K}(\xi\delta/2, \alpha)$ and $\widetilde{\mathbf{y}} \in R_{\widetilde{L},K}(\xi\delta/2, \alpha)$, from (3.22) in Lemma 3.5 we see that

$$|J_{\mathbf{y}}| = |y_{L+K+1} - y_{L-K-1}| = \frac{\mathcal{K}}{N\varrho_V(\bar{y})} + O(N^{-1}K^\xi) \tag{4.15}$$

$$|J_{\widetilde{\mathbf{y}}}| = |\widetilde{y}_{L+K+1} - \widetilde{y}_{L-K-1}| = \frac{\mathcal{K}}{N\varrho_G(\bar{\widetilde{y}})} + O(N^{-1}K^\xi). \tag{4.16}$$

Since ϱ_V is C^1, we have for any $|j| \le K$,

$$\begin{aligned}|\varrho_V(\bar{y}) - \varrho_V(\gamma_{L+j,V})| \le& C|\bar{y} - \gamma_{L+j,V}| \\ \le& C|\bar{y} - y_{L,V}| + C|\gamma_{L+j,V} - \gamma_{L,V}| + O(N^{-1}K^\xi) \le CKN^{-1},\end{aligned}$$

and similarly for $\varrho_G(\overline{\widetilde{y}})$.

Using (4.15), (4.14) and that the densities are separated away from zero, we easily obtain that

$$s := \frac{|J_{\mathbf{y}}|}{|J_{\widetilde{\mathbf{y}}}|} \qquad \text{satisfies} \qquad s = s_{\mathbf{y},\widetilde{\mathbf{y}}} = 1 + O(K^{-1+\xi}). \tag{4.17}$$

Note that this s is different from the scaling parameter in the first step but it will play a similar role so we use the same notation. For each fixed $\mathbf{y}, \widetilde{\mathbf{y}}$ we can now scale the conditional Gaussian log-gas $\mu_{\widetilde{\mathbf{y}}}$ by a factor s, i.e. change $\widetilde{\mathbf{y}}$ to $s\widetilde{\mathbf{y}}$, so that after rescaling $|J_{\mathbf{y}}| = |J_{s\widetilde{\mathbf{y}}}|$.

We will now show that this rescaling does not alter the gap statistics:

Lemma 4.3 *Suppose that s satisfies* (4.17) *and let $\mu = \mu_G$ be the Gaussian log-gas. Then we have*

$$\left|\left[\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}} - \mathbb{E}^{\mu_{s\widetilde{\mathbf{y}}}}\right]\mathcal{O}(\mathbf{x})\right| \le CK^{-1+\xi} \tag{4.18}$$

with

$$\mathcal{O}(\mathbf{x}) := O\big((N\varrho_m^G)(x_m - x_{m+1}), \ldots, (N\varrho_m^G)(x_m - x_{m+n})\big)$$

for any $\widetilde{L} - K \le m \le \widetilde{L} + K - n$ (note that the observable is not rescaled).

Proof. Define the Gaussian log-gas

$$\mu^s_{\widetilde{\mathbf{y}}} \sim e^{-N\beta\mathcal{H}^s_{\widetilde{\mathbf{y}}}}$$

with $\mathcal{H}^s_{\widetilde{\mathbf{y}}}$ defined exactly as in (3.6) but $V_{\mathbf{y}}(x)$ is replaced with

$$V^s_{\widetilde{\mathbf{y}}}(x) = V_s(x) - \frac{2}{N}\sum_{j\notin \widetilde{I}} \log|x - \widetilde{y}_j|, \qquad V_s(x) = \frac{1}{2}s^2x^2, \qquad \widetilde{I} := [\![\widetilde{L} - K, \widetilde{L} + K]\!].$$

Then by scaling

$$\mathbb{E}^{\mu_{s\widetilde{\mathbf{y}}}}\mathcal{O}(\mathbf{x}) = \mathbb{E}^{\mu^s_{\widetilde{\mathbf{y}}}}\mathcal{O}(\mathbf{x}/s) = \mathbb{E}^{\mu^s_{\widetilde{\mathbf{y}}}}\mathcal{O}(\mathbf{x}) + O(|s-1|), \tag{4.19}$$

where in the last step we used that the observable O is a smooth function with compact support. The error term is negligible by (4.17) and (3.1).

In order to control $\left[\mathbb{E}^{\mu^s_{\widetilde{\mathbf{y}}}} - \mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}\right]\mathcal{O}(\mathbf{x})$, it is sufficient to bound the relative entropy $S(\mu^s_{\widetilde{\mathbf{y}}}|\mu_{\widetilde{\mathbf{y}}})$. However, for any $\mathbf{y} \in R_{L,K}$ we have

$$\mathcal{H}''_{\mathbf{y}} \ge \min_{x\in J_{\mathbf{y}}} \frac{1}{N}\sum_{j\notin I}\frac{1}{|x-y_j|^2} \ge \frac{cN}{K} \tag{4.20}$$

with a positive constant. Applying this for $\widetilde{\mathbf{y}}$, we see that $\mu_{\widetilde{\mathbf{y}}}$ satisfies the logarithmic Sobolev inequality (LSI)

$$S(\mu^s_{\widetilde{\mathbf{y}}}|\mu_{\widetilde{\mathbf{y}}}) \le \frac{CK}{N} D(\mu^s_{\widetilde{\mathbf{y}}}|\mu_{\widetilde{\mathbf{y}}}),$$

where

$$S(\mu|\omega) := \int\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\omega}\log\frac{\mathrm{d}\mu}{\mathrm{d}\omega}\Big)\mathrm{d}\omega, \qquad D(\mu|\omega) := \frac{1}{N}\int\Big|\nabla\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\omega}}\Big|^2\mathrm{d}\omega$$

is the relative entropy and the relative Dirichlet form of two probability measures. Therefore

$$S(\mu^s_{\widetilde{\mathbf{y}}}|\mu_{\widetilde{\mathbf{y}}}) \le \frac{CK}{N^2}\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}\sum_{i\in\widetilde{I}}|NV'_s(x_i) - NV'(x_i)|^2 = CK(s^2-1)^2\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}\sum_{i\in\widetilde{I}}x_i^2 \le CK^4N^{-2}(s-1)^2.$$

In the last step we used (4.11) which, by rigidity for the Gaussian log-gas, guarantees that $|x_i| \leq CK/N$ with very high probability for any $i \in \widetilde{I}$. Together with (4.19) and (4.17) we obtain (4.18) and this proves Lemma 4.3. $\square$

Summarizing, we can from now on assume that (4.14) holds and that $\mathbf{y}, \widetilde{\mathbf{y}}$ satisfy $|J_{\mathbf{y}}| = |J_{\widetilde{\mathbf{y}}}|$. By a straightforward shift we can also assume that $J_{\mathbf{y}} = J_{\widetilde{\mathbf{y}}}$ so that the condition (3.8) of Theorem 3.1 is satisfied. The condition (3.9) has already been proved to hold in Lemma 3.5. Condition (3.10) follows from the definition of the sets $\mathcal{R}^*_{L,K,\mu}$ and $\mathcal{R}^*_{\widetilde{L},K,\widetilde{\mu}}$, see Lemma 4.1. Thus all conditions of Theorem 3.1 are verified. Finally, we remark that the multiplicative factors ϱ^V_{L+p} and $\varrho^G_{\widetilde{L}+p}$ in (4.13) coincide by (4.14) and (2.15). Then Theorem 3.1 (with an observable O rescaled by the common factor $\varrho^V_{L+p} = \varrho^G_{\widetilde{L}+p}$) implies Proposition 4.2. $\square$

5 Gap universality for Wigner matrices: proof of Theorem 2.2

In our recent results on the universality of Wigner matrices [21, 27, 28], we established the universality for Gaussian divisible matrices by establishing the local ergodicity of the Dyson Brownian motion (DBM). By local ergodicity we meant an effective estimate on the time to equilibrium for local average of observables depending on the gap. In fact, we gave an almost optimal estimate on this time. Then we used the Green function comparison theorem to connect Gaussian divisible matrices to general Wigner matrices. The local ergodicity of DBM was done by studying the flow of the global Dirichlet form. The estimate on the global Dirichlet form in all these works was sufficiently strong so that it implied the "ergodicity for locally averaged observables" without having to go through the local equilibrium measures. In the earlier work [23], however, we used an approach common in the hydrodynamical limits by studying the properties of local equilibrium measures. Since by Theorem 3.1 we now know the local equilibrium measures very well, we will now combine the virtue of both methods to prove Theorem 2.2. To explain the new method we will be using, we first recall the standard approach to the universality from [21, 27, 28] that consists of the following three steps:

i) rigidity estimates on the precise location of the eigenvalues.

ii) Dirichlet form estimates and local ergodicity of DBM.

iii) Green function comparison theorem to remove the small Gaussian convolution.

In order to prove the single gap universality, we will need to apply a similar strategy for the local equilibrium measure $\mu_{\mathbf{y}}$. However, apart from establishing rigidity for $\mu_{\mathbf{y}}$, we will need to strengthen Step ii). The idea is to use Dirichlet form estimates as in the previous approach, but we then use these estimates to show that the "local structure" after the evolution of the DBM for a short time is characterized by the local equilibrium $\mu_{\mathbf{y}}$ in a strong sense, i.e. without averaging. Since Theorem 3.1 provides a single gap universality for the local equilibrium $\mu_{\mathbf{y}}$, this proves the single gap universality after a short time DBM evolution and thus obtain the strong form of the Step ii) without averaging the observables. Notice that the key input here is Theorem 3.1 which contains an effective estimate on the time to equilibrium for each single gap. We will call this property the *strong local ergodicity of DBM.* In particular, our result shows that the local averaging taken in our previous works is not essential.

We now recall the rigidity estimate which asserts that the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ of a generalized Wigner matrix follow the Wigner semicircle law $\varrho_G(x)$ (2.6) in a very strong local sense. More precisely, Theorem 2.2. of [26] states that the eigenvalues are near their classical locations, $\{\gamma_j\}_{j=1}^N$, (2.7), in the sense that

$$\mathbb{P}\left\{\exists j \ : \ |\lambda_j - \gamma_j| \geq (\log N)^\zeta \Big[\min\big(j, N-j+1\big)\Big]^{-1/3} N^{-2/3}\right\} \leq C \exp\big[-c(\log N)^{\phi\zeta}\big] \tag{5.1}$$

for any exponent ζ satisfying

$$A_0 \log\log N \leq \zeta \leq \frac{\log(10N)}{10\log\log N}$$

where the positive constants C, ϕ, A_0, depend only on $C_{inf}, C_{sup}, \theta_1, \theta_2$, see (2.2), (2.4). In particular, for any fixed $\alpha > 0$ and $\nu > 0$, there are constants $C_0, c_1, c_2 > 0$ such that for any $N \geq 1$ and $k \in [\![\alpha N, (1-\alpha)N]\!]$, we have

$$\mathbb{P}\left(|\lambda_k - \gamma_k| > N^{-1+\nu}\right) \leq C_0 \exp\big(-c_1 N^{c_2}\big) \tag{5.2}$$

and (5.1) also implies

$$\mathbb{E}\sum_{k=1}^N (\lambda_k - \gamma_k)^2 \leq N^{-1+2\nu} \tag{5.3}$$

for any $\nu > 0$. The constants C_0, c_1, c_2 may be different from the ones in (4.1) but they play a similar role so we keep their notation. With a slight abuse of notation, we introduce the set $\mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi, \alpha)$ from (3.7) in the generalized Wigner setup as well, just γ_k denote the classical localitions with respect to the semicircle law, see (2.7). In particular (5.1) implies that for any $\xi, \alpha > 0$

$$\mathbb{P}\big(\mathcal{R}_{L,K}(\xi,\alpha)\big) \geq 1 - C_0 \exp\big(-c_1 N^{c_2}\big) \tag{5.4}$$

holds with some positive constants C_0, c_1, c_2, analogously to (4.4). We remark that the rigidity bound (5.1) for the generalized Wigner matrices is optimal throughout the spectrum and it gives a stronger control than the estimate used in the intermediate regime in the second line of the definition (3.7). For the forthcoming argument the weaker estimates are sufficient, so for notational simplicity we will not modify the definition of $\mathcal{R}$.

The Dyson Brownian motion (DBM) describes the evolution of the eigenvalues of a flow of Wigner matrices, $H = H_t$, if each matrix element h_{ij} evolves according to independent (up to symmetry restriction) Brownian motions. The dynamics of the matrix elements are given by an Ornstein-Uhlenbeck (OU) process which leaves the standard Gaussian distribution invariant. In the Hermitian case, the OU process for the rescaled matrix elements $v_{ij} := N^{1/2} h_{ij}$ is given by the stochastic differential equation

$$\mathrm{d}v_{ij} = \mathrm{d}\beta_{ij} - \frac{1}{2} v_{ij} \mathrm{d}t, \qquad i,j = 1,2,\ldots N, \tag{5.5}$$

where β_{ij}, $i < j$, are independent complex Brownian motions with variance one and β_{ii} are real Brownian motions of the same variance. The real symmetric case is analogous, just β_{ij} are real Brownian motions.

Denote the distribution of the eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$ of H_t at time t by $f_t(\boldsymbol{\lambda})\mu(\mathrm{d}\boldsymbol{\lambda})$ where the Gaussian measure μ is given by (2.5). The density $f_t = f_{t,N}$ satisfies the forward equation

$$\partial_t f_t = \mathcal{L} f_t, \tag{5.6}$$

where

$$\mathcal{L} = \mathcal{L}_N := \sum_{i=1}^N \frac{1}{2N}\partial_i^2 + \sum_{i=1}^N \Big(-\frac{\beta}{4}\lambda_i + \frac{\beta}{2N}\sum_{j\neq i}\frac{1}{\lambda_i - \lambda_j}\Big)\partial_i, \quad \partial_i = \frac{\partial}{\partial \lambda_i}, \tag{5.7}$$

with $\beta = 1$ for the real symmetric case and $\beta = 2$ in the complex hermitian case. The initial data f_0 given by the original generalized Wigner matrix.

Now we define a useful technical tool that was first introduced in [21]. For any $\tau > 0$ denote by $W = W^\tau$ an auxiliary potential defined by

$$W^\tau(\boldsymbol{\lambda}) := \sum_{j=1}^N W_j^\tau(\lambda_j), \qquad W_j^\tau(\lambda) := \frac{1}{2\tau}(\lambda_j - \gamma_j)^2, \tag{5.8}$$

i.e. it is a quadratic confinement on scale $\sqrt{\tau}$ for each eigenvalue near its classical location, where the parameter $\tau > 0$ will be chosen later.

Definition 5.1 *We define the probability measure* $\mathrm{d}\mu^\tau := Z_\tau^{-1} e^{-N\beta\mathcal{H}^\tau}$, *where the total Hamiltonian is given by*

$$\mathcal{H}^\tau := \mathcal{H} + W^\tau. \tag{5.9}$$

Here $\mathcal{H}$ *is the Gaussian Hamiltonian given by* (2.5) *and* $Z_\tau = Z_{\mu^\tau}$ *is the partition function. The measure* μ^τ *will be referred to as the relaxation measure with relaxation time* τ.

Denote by Q the following quantity

$$Q := \sup_{0\le t\le 1} \frac{1}{N}\int \sum_{j=1}^N (\lambda_j - \gamma_j)^2 f_t(\boldsymbol{\lambda})\mu(\mathrm{d}\boldsymbol{\lambda}). \tag{5.10}$$

Since H_t is a generalized Wigner matrix for all t, (5.3) implies that

$$Q \le N^{-2+2\nu} \tag{5.11}$$

for any $\nu > 0$ if $N \ge N_0(\nu)$ is large enough.

Recall the definition of the Dirichlet form w.r.t. a probability measure ω

$$D^\omega(\sqrt{g}) := \sum_{i=1}^N D_i^\omega(\sqrt{g}), \qquad D_i^\omega(\sqrt{g}) := \frac{1}{N}\int |\partial_i\sqrt{g}|^2 \mathrm{d}\omega = \frac{1}{4N}\int |\partial_i \log g|^2 g\,\mathrm{d}\omega,$$

and the definition of the relative entropy of two probability measures $g\omega$ and ω

$$S(g\omega|\omega) := \int g\log g\,\mathrm{d}\omega.$$

Now we recall Theorem 2.5 from [29]:

Theorem 5.2 *For any $\tau > 0$ and consider the local relaxation measure μ^τ. Set $\psi := \frac{d\mu^\tau}{d\mu}$ and let $g_t := f_t/\psi$. Suppose there is a constant m such that*

$$S(f_\tau \mu^\tau | \mu^\tau) \le CN^m. \tag{5.12}$$

Then for any $t \ge \tau N^{\varepsilon'}$ the entropy and the Dirichlet form satisfy the estimates:

$$S(f_t \mu^\tau | \mu^\tau) \le CN^2 Q \tau^{-1}, \qquad D^{\mu^\tau}(\sqrt{g_t}) \le CN^2 Q \tau^{-2}, \tag{5.13}$$

where the constants depend on ε' and m.

Corollary 5.3 *Fix $\mathfrak{a} > 0$ and let $\tau \ge N^{-\mathfrak{a}}$. Under the assumptions of Theorem 5.2, for any $t \ge \tau N^{\varepsilon'}$ the entropy and the Dirichlet form satisfy the estimates:*

$$D^\mu(\sqrt{f_t}) \le CN^2 Q \tau^{-2}. \tag{5.14}$$

Furthermore, if the initial data of the DBM, f_0, is given by a generalized Wigner ensemble, then

$$D^\mu(\sqrt{f_t}) \le CN^{2\mathfrak{a}+2\nu} \tag{5.15}$$

for any $\nu > 0$.

Proof. Since $g_t = f_t/\psi$, we have

$$\begin{aligned} D^\mu(\sqrt{f_t}) &= \sum_{i=1}^N \frac{1}{4N} \int |\partial_i \log g_t + \partial_i \log \psi|^2 f_t \mathrm{d}\mu \\ &\le \frac{1}{2N} \sum_{i=1}^N \int |\partial_i \log g_t|^2 f_t \mathrm{d}\mu + \frac{1}{2N} \sum_{i=1}^N \int |\partial_i \log \psi|^2 f_t \mathrm{d}\mu \\ &\le 2 \sum_{i=1}^N D^{\mu^\tau}(\sqrt{g_t}) + 2N^2 Q \tau^{-2}. \end{aligned}$$

Thus (5.14) follows from Theorem 5.2. Finally, (5.15) follows from (5.14) and (5.11). □

Define $f_{\mathbf{y}}$ to be the conditional density of $f\mu$ given $\mathbf{y}$ w.r.t. $\mu_{\mathbf{y}}$, i.e. it is defined by the relation $f_{\mathbf{y}}\mu_{\mathbf{y}} = (f\mu)_{\mathbf{y}}$. For any $\mathbf{y} \in \mathcal{R}_{L,K}$ we have the convexity bound (4.20). Thus we have the logarithmic Sobolev inequality

$$S(f_{\mathbf{y}}\mu_{\mathbf{y}} | \mu_{\mathbf{y}}) \le C \frac{K}{N} \sum_{i \in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) \tag{5.16}$$

and the bound

$$\int \mathrm{d}\mu_{\mathbf{y}} |f_{\mathbf{y}} - 1| \le \sqrt{S(f_{\mathbf{y}}\mu_{\mathbf{y}} | \mu_{\mathbf{y}})} \le C \sqrt{\frac{K}{N} \sum_{i \in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}})}. \tag{5.17}$$

To control the Dirichlet forms D_i for most external configurations $\mathbf{y}$, we need the following Lemma.

Lemma 5.4 *Fix $\mathfrak{a} > 0$, $\nu > 0$, and $\tau \geq N^{-\mathfrak{a}}$. Suppose the initial data f_0 of the DBM is given by a generalized Wigner ensemble. Then, with some small $\varepsilon' > 0$, for any $t \geq \tau N^{\varepsilon'}$ there exists a set $\mathcal{G}_{L,K} \subset \mathcal{R}_{L,K}$ of good boundary conditions $\mathbf{y}$ with*

$$\mathbb{P}^{f_t\mu}(\mathcal{G}_{L,K}) \geq 1 - CN^{-\varepsilon'}, \tag{5.18}$$

such that for any $\mathbf{y} \in \mathcal{G}_{L,K}$ we have

$$\sum_{i\in I} D_i^{\mu_\mathbf{y}}(\sqrt{f_{t,\mathbf{y}}}) \leq CN^{3\varepsilon'+2\mathfrak{a}+2\nu}, \quad f_{t,\mathbf{y}} = (f_t)_\mathbf{y}, \quad I = I_{L,K}, \tag{5.19}$$

and for any bounded observable O

$$\big|[\mathbb{E}^{f_{t,\mathbf{y}}\mu_\mathbf{y}} - \mathbb{E}^{\mu_\mathbf{y}}]O(\mathbf{x})\big| \leq CK^{1/2}N^{2\varepsilon'+\mathfrak{a}+\nu-1/2}. \tag{5.20}$$

Furthermore, for any $k \in I$ we also have

$$|\mathbb{E}^{f_{t,\mathbf{y}}\mu_\mathbf{y}} x_k - \gamma_k| \leq CN^{-1+\nu}. \tag{5.21}$$

Proof. In this proof, we omit the subscript t, i.e. we use $f = f_t$. By definition of the conditional measure and by (5.15), we have

$$\mathbb{E}^{f\mu}\sum_{i\in I} D_i^{\mu_\mathbf{y}}(\sqrt{f_\mathbf{y}}) = \sum_{i\in I} D_i^{\mu}(\sqrt{f}) \leq D^\mu(\sqrt{f}) \leq CN^{4\mathfrak{a}+2\nu}.$$

By Markov inequality, (5.19) holds for all $\mathbf{y}$ in a set $\mathcal{G}^1_{L,K}$ with $\mathbb{P}^{f\mu}(\mathcal{G}^1_{L,K}) \geq 1 - CN^{-3\varepsilon'}$. Without loss of generality, by (5.4) we can assume that $\mathcal{G}^1_{L,K} \subset \mathcal{R}_{L,K}$. The estimate (5.20) now follows from (5.19) and (5.16).

Similarly, the rigidity bound (5.2) with respect to $f\mu$ can be translated to the measure $f_\mathbf{y}\mu_\mathbf{y}$, i.e. there exists a set $\mathcal{G}^2_{L,K}$, with

$$\mathbb{P}^{f\mu}(\mathcal{G}^2_{L,K}) \geq 1 - C_0 \exp\big(-\frac{1}{2}c_1 N^{c_2}\big),$$

such that for any $\mathbf{y} \in \mathcal{G}^2_{L,K}$ and for any $k \in I$, we have

$$\mathbb{P}^{f_\mathbf{y}\mu_\mathbf{y}}\Big(|x_k - \gamma_k| \geq N^{-1+\nu}\Big) \leq \exp\big(-\frac{1}{2}c_1 N^{c_2}\big).$$

In particular, we can conclude (5.21) for any $\mathbf{y} \in \mathcal{G}^2_{L,K}$. Setting $\mathcal{G}_{L,K} := \mathcal{G}^1_{L,K} \cap \mathcal{G}^2_{L,K}$ we proved the lemma. $\square$

Lemma 5.5 *Fix $\mathfrak{a} > 0$, $\nu > 0$, and $\tau \geq N^{-\mathfrak{a}}$. Suppose the initial data f_0 of the DBM is given by a generalized Wigner ensemble. Then, with some small $\varepsilon' > 0$, for any $t \geq \tau N^{\varepsilon'}$, $k \in I$ and $\mathbf{y} \in \mathcal{G}_{L,K}$, we have*

$$|\mathbb{E}^{\mu_\mathbf{y}} x_k - \mathbb{E}^{f_{t,\mathbf{y}}\mu_\mathbf{y}} x_k| \leq KN^{-3/2+\nu+\mathfrak{a}+2\varepsilon'}. \tag{5.22}$$

In particular, if the parameters chosen such that

$$KN^{-3/2+\nu+\mathfrak{a}+2\varepsilon'} \leq N^{-1}K^\xi, \qquad \text{and} \quad N^{-1+\nu} \leq N^{-1}K^\xi$$

with some small $\xi > 0$, then

$$\big|\mathbb{E}^{\mu_\mathbf{y}} x_k - \alpha_k\big| \leq CN^{-1}K^\xi, \qquad k \in I, \tag{5.23}$$

where α_k is defined in (3.4). In other words, the analogue of (3.10) is satisfied.

Notice that if we apply (5.20) with the special choice $O(\mathbf{x}) = x_k$ then the error estimate would be much worse than (5.22). We wish to emphasize that (5.23) is not an obvious fact although we know that it holds for $\mathbf{y}$ with high probability w.r.t. the equilibrium measure μ. The key point of (5.23) is that it holds for any $\mathbf{y} \in \mathcal{G}_{L,K}$ and thus with "high probability" w.r.t $f_t\mu$!

Proof. Once again, we omit the subscript t. The estimate (5.23) is a simple consequence of (5.22), (5.21) and (4.10). To prove (5.22), we run the reversible dynamics

$$\partial_s q_s = \mathcal{L}_{\mathbf{y}} q_s \tag{5.24}$$

starting from initial data $q_0 = f_{\mathbf{y}}$, where the generator $\mathcal{L}_{\mathbf{y}}$ is the unique reversible generator with the Dirichlet form $D^{\mu_{\mathbf{y}}}$, i.e.,

$$-\int f \mathcal{L}_{\mathbf{y}}\, g \,\mathrm{d}\mu_{\mathbf{y}} = \sum_{i\in I} \frac{1}{N} \int \nabla_i f \cdot \nabla_i g \,\mathrm{d}\mu_{\mathbf{y}}.$$

Recall that from the convexity bound (4.20), $\tau_K = K/N$ is the time to equilibrium of this dynamics. After differentiation and integration we get,

$$\Big| \mathbb{E}^{\mu_{\mathbf{y}}} x_k - \mathbb{E}^{f_{\mathbf{y}}\mu_{\mathbf{y}}} x_k \Big| = \Big| \int_0^{K^{\varepsilon'}\tau_K} \mathrm{d}u \frac{1}{N} \int (\partial_k q_u) \mathrm{d}\mu_{\mathbf{y}} \Big| + O(\exp{(-cK^{\varepsilon'})}).$$

From the Schwarz inequality with a free parameter R, we can bound the last line by

$$\frac{1}{N} \int_0^{K^{\varepsilon'}\tau_K} \mathrm{d}u \int \Big(R(\partial_k \sqrt{q_u})^2 + R^{-1} \Big) \mathrm{d}\mu_{\mathbf{y}} + O(\exp{(-cK^{\varepsilon'})}).$$

Dropping the trivial subexponential error term and using that the time integral of the Dirichlet form is bounded by the initial entropy, we can bound the last line by

$$RS(f_{\mathbf{y}}\mu_{\mathbf{y}} | \mu_{\mathbf{y}}) + \frac{K^{\varepsilon'}\tau_K}{NR}.$$

Using the logarithmic Sobolev inequality for $\mu_{\mathbf{y}}$ and optimizing the parameter R, we can bound the last term by

$$\begin{aligned} \Big| \mathbb{E}^{\mu_{\mathbf{y}}} x_k - \mathbb{E}^{f_{\mathbf{y}}\mu_{\mathbf{y}}} x_k \Big| &\le \tau_K R \sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) + \frac{K^{\varepsilon'}\tau_K}{NR} + O(\exp{(-cK^{\varepsilon'})}) \\ &\le \frac{K^{\varepsilon'}\tau_K}{N^{1/2}} \Big(\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) \Big)^{1/2} + O(\exp{(-cK^{\varepsilon'})}). \end{aligned} \tag{5.25}$$

Combining this bound with (5.19), we obtain (5.22). □

We now prove the following comparison for the local statistics of μ and $f_t\mu$, where μ is the Gaussian β-ensemble, (2.11), with quadratic V, and f_t is the solution of (5.6) with initial data f_0 given by the original generalized Wigner matrix.

Lemma 5.6 *Fix $n > 0$, $\mathfrak{a} > 0$ and $\tau \ge N^{-\mathfrak{a}}$. Then for sufficient small $\mathfrak{a}$ there exist positive ε and ε' such that for any $t \ge \tau N^{\varepsilon'}$, for any n and for any compactly supported smooth observable O we have*

$$\Big| \big[\mathbb{E}^{f_t\mu} - \mathbb{E}^{\mu}\big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n})\big) \Big| \le CN^{-\varepsilon}, \tag{5.26}$$

for any $j \in [\![\alpha N, (1-\alpha)N]\!]$ and for any sufficiently large N.

Proof. We will apply Lemma 5.4, and we choose $L = j$. Since $K \leq N^{1/4}$, the right hand side of (5.20) is smaller than $N^{-\varepsilon}$. Then we have

$$\Big| \big[\mathbb{E}^{f_{t,\mathbf{y}} \mu_{\mathbf{y}}} - \mathbb{E}^{\mu_{\mathbf{y}}} \big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n}) \big) \Big| \leq C N^{-\varepsilon}, \tag{5.27}$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$ with the probability of $\mathcal{G}_{L,K}$ satisfying (5.18). Choose any $\tilde{\mathbf{y}} \in \mathcal{R}^*$ defined in Lemma 4.1. We now apply Theorem 3.1 with both $\mu_{\mathbf{y}}$ and $\mu_{\tilde{\mathbf{y}}}$ given by local Gaussian β-ensemble. Thus the estimate (3.10) is guaranteed by (4.6) and (5.23). Since $\mathbf{y}, \tilde{\mathbf{y}} \in \mathcal{R} = \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$ and Lemma 3.5 guarantees (3.9), we can apply Theorem 3.1 so that

$$\Big| \big[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\mu_{\tilde{\mathbf{y}}}} \big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n}) \big) \Big| \leq C N^{-\varepsilon}, \tag{5.28}$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$ and $\tilde{\mathbf{y}} \in \mathcal{R}^*$. Since $\mathbb{P}^{\mu}(\mathcal{R}^*) \geq 1 - N^{-\varepsilon}$, see (4.5), we have thus proved that

$$\Big| \big[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\mu} \big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n}) \big) \Big| \leq C N^{-\varepsilon}, \tag{5.29}$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$. From (5.27), (5.29) and the probability estimate (5.18) for $\mathcal{G}_{L,K}$, with possibly reducing ε so that $\varepsilon \leq \varepsilon'$, we obtain that

$$\Big| \big[\mathbb{E}^{f_t \mu} - \mathbb{E}^{\mu} \big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n}) \big) \Big| \leq C N^{-\varepsilon}. \tag{5.30}$$

This proves Lemma 5.6. □

Recall that H_t is the generalized Wigner matrix whose matrix elements evolve by independent OU processes. Thus in Lemma 5.6 we have proved that the local statistics of H_t, for $t \geq N^{-2\mathfrak{a}+\varepsilon'}$, is the same as the corresponding Gaussian one for any initial generalized matrix H_0. Finally, we need to approximate the generalized Wigner ensembles by Gaussian divisible ones. The idea of approximation first appeared in [19] via a "reverse heat flow" argument and was also used in [46] via a four moment theorem. We will follow the Green function comparison theorem of [27, 28] and in particular, the result in [37] since these results were formulated and proved for the generalized Wigner matrices.

Theorem 1.10 from [37] implies that if the first four moments of two generalized Wigner ensembles, $H^{\mathbf{v}}$ and $H^{\mathbf{w}}$, are the same and a certain level repulsion estimate holds for one of the ensembles, say, $H^{\mathbf{v}}$, then we have

$$\lim_{N \to \infty} \big[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}} \big] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \ldots, N(x_j - x_{j+n}) \big) \; = \; 0. \tag{5.31}$$

In [37] it was assumed that one of the ensemble satisfies a level repulsion estimate. This estimate, however, was needed only for eigenvector universality; following the argument in [37] one can check that no level repulsion was actually used for the eigenvalue universality. Roughly speaking, the proof of [37] was based on the Green function comparison theorem [27] which compares eigenvalues at a fixed energy. In order to convert fixed energy to a fixed eigenvalue index, one needs to know that the total number of eigenvalues up to a fixed energy is the same for the two ensembles. The total number of eigenvalues up to a fixed energy E can be expressed in terms of integration of imaginary part of the trace of Green functions, i.e.,

$$\int_{-\infty}^{E} \mathrm{d}y \, \mathrm{Im} \, \mathrm{Tr} \, \frac{1}{H - (y + i\eta)}$$

with an η slightly smaller than $1/N$. Thus the basic idea of the Green function comparison theorem can be employed and this leads to (5.31). As in the Green function comparison theorem, the matching of the fourth moments can be relaxed to that the differences of the fourth moments are of order N^{-c} for any $c > 0$.

We now apply (5.31) with the choice $H^{\mathbf{v}}$ being the generalized Wigner ensemble for which we wish to prove the universality and $H^{\mathbf{w}} = H_t$ with $t = N^{-c'}$ for some small c'. Following [28], we construct an auxiliary Wigner matrix H_0 such that the first three moments of H_t and the *original* matrix $H^{\mathbf{v}}$ are identical while the differences of the fourth moments are less than $N^{-c''}$ for some small $c'' > 0$ depending on c' (see Lemma 3.4 of [28]). This completes the proof of (2.8) showing that the local gap statistics with the same gap-label (j) is identical for the Wigner matrix and the Gaussian case. The proof of (2.9) follows now directly from Theorem 2.3 that, in particular, compares the local gap statistics for different gap labels (k and m) in the Gaussian case. This completes the proof Theorem 2.2. $\square$

6 Rigidity and level repulsion of local measures

6.1 Rigidity of $\mu_{\mathbf{y}}$: proof of Theorem 3.2

We will prove Theorem 3.2 using a method similar to the proof of Theorem 3.1 in [7]. Theorem 3.1 of [7] was proved by a quite complicated argument involving induction on scales and the loop equation. The loop equation, however, requires analyticity of the potential and it cannot be applied to prove Theorem 3.2 for a local measure whose potential $V_{\mathbf{y}}$ is not analytic. We note, however, that in [7] the loop equation was used only to estimate the *expected locations* of the particles. Now this estimate is given as a condition by (3.12) and thus we can adapt the proof in [7] to the current setting. For later application, however, we will need a stronger form of the rigidity bound, namely we will establish that the tail of the gap distribution has a Gaussian decay. This stronger statement requires some modifications to the argument from [7] which therefore we partially repeat here. We now introduce the notations needed to prove Theorem 3.2.

Let θ be a continuously differentiable nonnegative function with $\theta = 0$ on $[-1, 1]$ and $\theta'' \geq 1$ for $|x| > 1$. We can take for example $\theta(x) = (x-1)^2 \mathbb{1}_{x>1} + (x+1)^2 \mathbb{1}_{x<-1}$ in the following.

For any $m \in [\![\alpha N, (1-\alpha)N]\!]$ and any integer $1 \leq M \leq \alpha N$, we denote $I^{(m,M)} = [\![m-M, m+M]\!]$ and $\mathcal{M} = |I^{(m,M)}| = 2M+1$. Let $\eta := \xi/3$. For any k, M with $|k - L| \leq K - M$, define

$$\phi^{(k,M)}(\mathbf{x}) := \sum_{i<j,\; i,j \in I^{(k,M)}} \theta\left(\frac{N(x_i - x_j)}{\mathcal{M}K^{2\eta}}\right). \tag{6.1}$$

Let

$$\omega_{\mathbf{y}}^{(k,M)} = Z_{\mathbf{y},\phi} \mu_{\mathbf{y}} e^{-\phi^{(k,M)}},$$

where $Z_{\mathbf{y},\phi}$ is a normalization constant. Choose an increasing sequence of integers, $M_1 < M_2 < \ldots < M_A$ such that $M_1 = K^{\xi}$, $M_A = CK^{1-2\eta}$ with a large constant C, and $M_\gamma / M_{\gamma-1} \sim K^{\eta}$ (meaning that $cK^{\eta} \leq M_\gamma/M_{\gamma-1} \leq CK^{\eta}$). We can choose the sequence such that $A \leq C\eta^{-1}$. We set $\omega_\gamma := \omega_{\mathbf{y}}^{(k,M_\gamma)}$ and we study the rigidity properties of the measures $\omega_A, \omega_{A-1}, \ldots, \omega_1$ in this order. Note that $\mu_{\mathbf{y}} = \omega_A$ since $\mathbf{y} \in \mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$ guarantees that $|x_i - x_j| \leq |J_{\mathbf{y}}| \leq CK/N$, see (3.22), thus for $M = M_A = CK^{1-2\eta}$ the argument of θ in (6.1) is smaller than 1, so

$\phi \equiv 0$ in this case. We also introduce the notation

$$x_k^{[M]} := \frac{1}{2M+1} \sum_{j=k-M}^{k+M} x_j .$$

Definition 6.1 *We say that $\mu_{\mathbf{y}}$ has exponential rigidity on scale ℓ if there are constants C, c, such that the following bound holds*

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq \ell + uK^{\xi}N^{-1}) \leq Ce^{-cu^2}, \qquad u > 0, \tag{6.2}$$

for any $k \in I$.

First we prove that $\mu_{\mathbf{y}}$ has exponential rigidity on scale $M_A N^{-1}$. Starting from $\gamma = A$, by the Herbst bound and the logarithmic Sobolev inequality for $\mu_{\mathbf{y}}$ with LSI constant of order K/N (5.16), we have for any $k \in [\![L - K + M_A, L + K - M_A]\!]$ that

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_k^{[M_A]} - \mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]}\right| \geq \frac{b}{\sqrt{M_A}}\right) \leq e^{-c(N/K)Nb^2}, \qquad b \geq 0, \tag{6.3}$$

i.e.

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_k^{[M_A]} - \mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]}\right| \geq \frac{uK^{\eta}}{N}\right) \leq Ce^{-cu^2}. \tag{6.4}$$

Using the estimate (5.23) we have that

$$\left|\mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]} - \alpha_k^{[M_A]}\right| \leq CN^{-1}K^{\xi}.$$

Thus we obtain

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_k^{[M_A]} - \alpha_k^{[M_A]}\right| \geq CN^{-1}K^{\xi} + \frac{uK^{\eta}}{N}\right) \leq Ce^{-cu^2}. \tag{6.5}$$

Since $x_{k-M}^{[M]} \leq x_k \leq x_{k+M}^{[M]}$ and the α_k's are regular with spacing of order $1/N$, we get

$$x_k - \alpha_k \leq x_{k+M}^{[M]} - \alpha_{k-M}^{[M]} \leq x_{k+M}^{[M]} - \alpha_{k+M}^{[M]} + CMN^{-1}$$

and we also have a similar lower bound. Thus

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_k - \alpha_k\right| \geq CM_A N^{-1} + \frac{uK^{\eta}}{N}\right) \leq Ce^{-cu^2} \tag{6.6}$$

for any $k \in [\![L-K+2M_A, L+K-2M_A]\!]$, where we used that $M_A \geq K^{\xi}$. If $k \in [\![L-K, L-K+2M_A]\!]$, then

$$x_k - \alpha_k \leq x_{L-K+2M_A} - \alpha_{L-K+2M_A} + CM_A N^{-1}$$

and

$$x_k - \alpha_k \geq y_{L-K-1} - \alpha_k \geq -CM_A N^{-1}.$$

Thus we have the estimate

$$|x_k - \alpha_k| \leq |x_{L-K+2M_A} - \alpha_{L-K+2M_A}| + CM_A N^{-1}.$$

Since (6.6) holds for the difference $x_{L-K+2M_A} - \alpha_{L-K+2M_A}$, we have that it holds for $x_k - \alpha_k$ as well (with at most an adjustment of C) for any $k \in [\![L - K, L - K + 2M_A]\!]$. Similar argument holds for $k \in [\![L + K - 2M_A, L + K]\!]$. Thus we proved (6.6) for all $k \in [\![L - K, L + K]\!]$, i.e. we showed exponential rigidity on scale $M_A N^{-1}$.

Now we use an induction on scales and we show that if

(i) for any $k \in [\![L-K+M_\gamma, L+K-M_\gamma]\!]$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(|x_k^{[M_\gamma]} - \alpha_k^{[M_\gamma]}| \geq uK^\xi N^{-1}\right) \leq Ce^{-cu^2}, \qquad u \geq 0; \tag{6.7}$$

(ii) exponential rigidity holds on some scale $M_\gamma N^{-1}$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(|x_k - \alpha_k| \geq CM_\gamma N^{-1} + uK^\xi N^{-1}\right) \leq Ce^{-cu^2}, \qquad k \in I, \quad u \geq 0; \tag{6.8}$$

(iii) we have the entropy bound

$$S(\mu_{\mathbf{y}}|\omega_\gamma) \leq Ce^{-cM_\gamma^2 K^{-5\eta}}, \tag{6.9}$$

then (i)–(iii) also hold with γ replaced by $\gamma - 1$ as long as $M_{\gamma-1} \geq K^\xi$. The iteration can be started from $\gamma = A$, since (6.7) and (6.8) were proven in (6.5) and in (6.6) (even with a better bound), and (6.9) is trivial for $\gamma = A$ since $\omega_A = \mu_{\mathbf{y}}$.

We first notice that on any scale M_γ, the bound (6.7) implies (6.8) by the same argument as we concluded (6.6) for any $k \in I$ from (6.5). So we can focus on proving (6.7) and (6.9) on the scale $M_{\gamma-1}$.

To prove (6.9) on scale $M_{\gamma-1}$, notice that (6.8), with the choice $u = M_\gamma K^{-\xi}$, implies

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq CM_\gamma N^{-1}) \leq Ce^{-cM_\gamma^2 K^{-2\xi}}, \qquad k \in I. \tag{6.10}$$

Since

$$\theta\left(\frac{N(x_i - x_j)}{\mathcal{M}_{\gamma-1}K^{2\eta}}\right) = 0$$

unless $|x_i - x_j| \geq CM_{\gamma-1}N^{-1}K^{2\eta} \geq CM_\gamma N^{-1}K^\eta$, we have that the scale $CM_\gamma N^{-1}$ is by a factor K^η smaller than the scale of $x_i - x_j$ built into the definition of $\phi^{(k,M_{\gamma-1})}$, see (6.1). But for $i, j \in I^{(k,M_{\gamma-1})}$ we have $|x_i - x_j| \leq |x_i - \alpha_i| + |x_j - \alpha_j| + CM_{\gamma-1}N^{-1}$. Thus $\phi^{(k,M_{\gamma-1})} = 0$ unless we are on the event described in (6.10) at least for one k. Moreover, $|\nabla\phi^{(k,M_{\gamma-1})}(\mathbf{x})| \leq N^C$ for any configuration $\mathbf{x}$ in J. Thus, following the argument in Lemma 3.15 of [7], via the logarithmic Sobolev inequality for $\mu_{\mathbf{y}}$, we get

$$S(\mu_{\mathbf{y}}|\omega_{\gamma-1}) \leq CKN^{-1}\mathbb{E}^{\mu_{\mathbf{y}}}|\nabla\phi^{(k,M_{\gamma-1})}|^2 \leq CN^C e^{-cM_\gamma^2 K^{-2\xi}} \leq Ce^{-cM_{\gamma-1}^2 K^{-5\eta}}. \tag{6.11}$$

Here we used that the prefactor N^C can be absorbed in the exponent by using that $M_\gamma^2 K^{-2\xi} - M_{\gamma-1}^2 K^{-5\eta} \geq K^{2\xi - 5\eta} = K^\eta \geq N^{\eta\delta}$. Here we have used $\xi = 3\eta$ and $M_{\gamma-1} \geq K^\xi$. We will not need it here, but we note that the same bound on the opposite relative entropy,

$$S(\omega_{\gamma-1}|\mu_{\mathbf{y}}) \leq Ce^{-cM_{\gamma-1}^2 K^{-5\eta}},$$

is also correct. Thus (6.9) for $\gamma - 1$ is proved.

Now we focus on proving (6.7) on the scale $M_{\gamma-1}$. Set $1 \leq M' \leq M \leq K$ and fix an index $k \in I$ such that $|k - L| \leq K - M$. We state the following slightly generalized version of Lemma 3.14 of [7]

Lemma 6.2 *For any integers $1 \leq M' \leq M \leq K$, $k \in [\![L-K+M, L+K-M]\!]$ and $k' \in [\![k-M+M', k+M-M']\!]$, we have*

$$\mathbb{P}^{\omega^{(k,M)}}\left(\left|\lambda_{k'}^{[M']} - \lambda_k^{[M]} - \mathbb{E}^{\omega^{(k,M)}}\left(\lambda_{k'}^{[M']} - \lambda_k^{[M]}\right)\right| > \frac{uK^{2\eta}}{N}\sqrt{\frac{M}{M'}}\right) \leq Ce^{-cu^2}.$$

Compared with Lemma 3.14 of [7], we first note that N^ε in Lemma 3.14 [7] is changed to $K^{2\eta}$ due to that $\phi^{(k,M)}(\mathbf{x})$ in (6.1) is defined with a $K^{2\eta}$ factor instead of N^ε. Furthermore, here we allowed the center at the scale M' to be different from k. The only condition is that the interval $[\![k'-M',k'+M']\!] \subset [\![k-M,k+M]\!]$. The proof of this lemma is identical to that of Lemma 3.14 of [7].

In particular, for any $\gamma = 2,3,\dots A$ and with $M' = M_{\gamma-1}$ and $M = M_\gamma \le K^\eta M_{\gamma-1}$ and with any choice of $k_\gamma \in [\![L-K+M_\gamma, L+K-M_\gamma]\!]$, $k_{\gamma-1} \in [\![L-K+M_{\gamma-1}, L+K-M_{\gamma-1}]\!]$, so that $[\![k_{\gamma-1}-M_{\gamma-1}, k_{\gamma-1}+M_{\gamma-1}]\!] \subset [\![k_\gamma - M_\gamma, k_\gamma + M_\gamma]\!]$, we get

$$\mathbb{P}^{\omega_\gamma}\left(\left|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]} - \mathbb{E}^{\omega_\gamma}\left(x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]}\right)\right| > \frac{uK^{5\eta/2}}{N}\right) \le Ce^{-cu^2}. \tag{6.12}$$

The entropy bound (6.9) and the boundedness of x_k imply that

$$\left|\mathbb{E}^{\omega_\gamma} x_k - \mathbb{E}^{\mu_{\mathbf{y}}} x_k\right| \le C\sqrt{S(\mu_{\mathbf{y}}|\omega_\gamma)} \le Ce^{-cM_\gamma^2 K^{-5\eta}};$$

where $M_\gamma^2 K^{-5\eta} \ge K^{2\xi-5\eta} \ge K^\eta$ $(\eta = \xi/3)$. We can combine it with (3.12) to have

$$\left|\mathbb{E}^{\omega_\gamma} x_k - \alpha_k\right| \le CK^\xi/N.$$

The measure ω_γ in (6.12) can also be changed to $\mu_{\mathbf{y}}$ at the expense of an entropy term $S(\mu_{\mathbf{y}}|\omega_\gamma)$. Using (6.9), we thus have

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]} - \left(\alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_\gamma}^{[M_\gamma]}\right)\right| \ge CK^\xi N^{-1} + \frac{uK^{5\eta/2}}{N}\right) \le Ce^{-cu^2} + Ce^{-cM_\gamma^2 K^{-5\eta}}. \tag{6.13}$$

Combining it with (6.7) and recalling $\xi = 3\eta$, we get

$$\mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}\right| \ge CK^\xi N^{-1} + \frac{uK^\xi}{N}\right) \le Ce^{-cu^2} + Ce^{-cM_\gamma^2 K^{-5\eta}}. \tag{6.14}$$

This gives (6.7) on scale $M_{\gamma-1}$ if $u \le cM_\gamma K^{-5\eta/2}$ with a small constant c. Suppose now that $u \ge cM_\gamma K^{-5\eta/2}$, which, in particular, means that $u \ge cK^{-\eta/2}$. Then, by (6.8), we have

$$\begin{aligned}
\mathbb{P}^{\mu_{\mathbf{y}}}\Big(\left|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}\right| &\ge CK^\xi N^{-1} + \frac{uK^\xi}{N}\Big) \\
&\le \mathbb{P}^{\mu_{\mathbf{y}}}\left(\left|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}\right| \ge CM_\gamma N^{-1} + (1-CK^{-\eta/2})u\frac{K^\xi}{N}\right) \\
&\le \sum_{k\in I} \mathbb{P}^{\mu_{\mathbf{y}}}\left(|x_k - \alpha_k| \ge CM_\gamma N^{-1} + (1-CK^{-\eta/2})u\frac{K^\xi}{N}\right) \\
&\le CKe^{-c(1-CK^{-\eta/2})^2u^2} \le Ce^{-c'u^2}.
\end{aligned}$$

This proves (6.7) for $\gamma - 1$. Note that the constants slightly deterioriate at each iteration step, but the number of iterations is finite (of order $1/\eta = 3/\xi$), so eventually the constants C, c in (3.13) may depend on ξ. In fact, since the deterioriation is minor, one can also prove (3.13) with ξ-independent constants, but for simplicity of the presentation we did not follow the change of these constants at each step.

After completing the iteration, from (6.8) for $\gamma = 1$, $M_1 = K^\xi$, we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}\Big(|x_k - \alpha_k| \geq CK^\xi N^{-1} + uK^\xi N^{-1}\Big) \leq Ce^{-cu^2}, \qquad k \in I;$$

This concludes (3.13) for $u \geq 1$. Finally, (3.13) is trivial for $u \leq 1$ if the constant C is sufficiently large. This completes the proof of Theorem 3.2. $\square$

6.2 Level repulsion estimates of $\mu_{\mathbf{y}}$: proof of Theorem 3.3

We now prove the level repulsion estimate, Theorem 3.3, for the local log-gas $\mu_{\mathbf{y}}$ with good boundary conditions $\mathbf{y}$. There are two key ideas in the following argument. We first recall the weak level repulsion estimate (4.11) in [7], which in the current notation asserts

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq CNs$$

for any $s > 0$, and similar estimates may be deduced for internal gaps. Compared with (3.14), this estimate does not contain any β exponent, moreover, in order to obtain (3.17), the N factor has to be reduced to K^ξ (neglecting the irrelevant $\log N$ factor). Our first idea is to run this proof for a local measure with only K^ξ particles to reduce the N factor to K^ξ. The second idea involves introducing some auxiliary measures to catch some of the β related factors. We first introduce these two auxiliary measures which are slightly modified versions of the local equilibrium measures:

$$\mu_0 := \mu_{\mathbf{y},0} = Z_0(x_{L-K} - y_{L-K-1})^{-\beta}\mu_{\mathbf{y}}; \quad \mu_1 := \mu_{\mathbf{y},1} = Z_1 W^{-\beta}\mu_{\mathbf{y}}, \tag{6.15}$$

$$W = (x_{L-K} - y_{L-K-1})(x_{L-K+1} - y_{L-K-1}), \tag{6.16}$$

where Z_0, Z_1 are chosen for normalization. In other words, we drop the term $(x_{L-K} - y_{L-K-1})^\beta$ from the measure $\mu_{\mathbf{y}}$ in μ_0 and we drop W^β in μ_1. To estimate the upper gap, $y_{L+K+1} - x_{L+K}$, similar results will be needed when we drop the term $(y_{L+K+1} - x_{L+K})^\beta$ and the analogous version of W, but we will not state them explicitly. We first prove the following results which are weaker than Theorem 3.3.

Lemma 6.3 *Let L and K satisfy* (3.1) *and consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in* (3.5)*.*

i) Let ξ, α be any fixed positive constants and let $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Then for any $s > 0$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{L-K} - y_{L-K-1} \leq s/N] \leq C\,(Ks\log N)^{\beta+1}, \tag{6.17}$$

and

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{L-K+1} - y_{L-K-1} \leq s/N] \leq C\,(Ks\log N)^{2\beta+1}. \tag{6.18}$$

ii) Let $\mathbf{y}$ be arbitrary with the only condition that $|y_i| \leq C$ for all i. Then for any $s > 0$ we have the weaker estimate

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{L-K} - y_{L-K-1} \leq s/N] \leq \left(\frac{CsK}{|J_{\mathbf{y}}|}\right)^{\beta+1}, \tag{6.19}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}[x_{L-K+1} - y_{L-K-1} \leq s/N] \leq \left(\frac{CsK}{|J_{\mathbf{y}}|}\right)^{2\beta+1}, \qquad j = 0, 1. \tag{6.20}$$

To prove Lemma 6.3, we first prove estimates even weaker than (6.17)–(6.20) for $\mu_{\mathbf{y}}$ and $\mu_{\mathbf{y},j}$.

Lemma 6.4 *Let L and K satisfy (3.1).*

i) Let ξ, α be any fixed positive constants and let $\mathbf{y} \in \mathcal{R}_{L.K} = \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$, then we have for any $s > 0$

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq CKs \log N, \tag{6.21}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq CKs \log N, \qquad j = 0, 1. \tag{6.22}$$

ii) Let $\mathbf{y}$ be arbitrary with the only condition that $|y_i| \leq C$ for all i. Then for any $s > 0$ we have the weaker estimate

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq \frac{CsK}{|J_{\mathbf{y}}|}, \tag{6.23}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq \frac{CsK}{|J_{\mathbf{y}}|}, \qquad j = 0, 1. \tag{6.24}$$

Proof. We will prove (6.21), the same proof with only change of notations works for (6.22) case as well. We will comment on this at the end of the proof.

For notational simplicity, we first shift the coordinates by S such that in the new coordinates $\bar{y} = 0$, i.e. $y_{L-K-1} = -y_{L+K+1}$ and J is symmetric to the origin. With the notation $a := -y_{L-K-1}$ and $I = [\![L-K, L+K]\!]$, we first estimate the following quantity, for any $0 \leq \varphi \leq c$ (with a small constant)

$$\begin{aligned} Z_\varphi &:= \int \ldots \int_{-a+a\varphi}^{a-a\varphi} \mathrm{d}\mathbf{x} \prod_{\substack{i,j \in I \\ i<j}} (x_i - x_j)^\beta e^{-N\frac{\beta}{2}\sum_j V_{\mathbf{y}}(S+x_j)} \\ &= (1-\varphi)^{K+\beta K(K-1)/2} \int \ldots \int_{-a}^{a} \mathrm{d}\mathbf{w} \prod_{i<j} (w_i - w_j)^\beta e^{-N\frac{\beta}{2}\sum_j V_{\mathbf{y}}(S+(1-\varphi)w_j)}, \end{aligned}$$

where we set

$$w_j := (1-\varphi)^{-1} x_{L+j}, \qquad \mathrm{d}\mathbf{x} = \prod_{|j|\leq K} \mathrm{d}x_{L+j} \qquad \mathrm{d}\mathbf{w} = \prod_{|j|\leq K} \mathrm{d}w_j. \tag{6.25}$$

By definition,

$$e^{-N\frac{\beta}{2}V_{\mathbf{y}}(S+(1-\varphi)w_j)} = e^{-N\frac{\beta}{2}V(S+(1-\varphi)w_j)} \prod_{k\leq L-K-1} ((1-\varphi)w_j - y_k)^\beta \prod_{k \geq L+K+1} (y_k - (1-\varphi)w_j)^\beta. \tag{6.26}$$

For the smooth potential V, we have

$$\Big| V(S + (1-\varphi)w_j)) - V(S + w_j) \Big| \leq C|\varphi w_j| \leq \frac{CK\varphi}{N} \tag{6.27}$$

with a constant depending on V, where we have used $|w_j| \leq a \leq CK/N$ which follows from $|J_{\mathbf{y}}| \leq CK/N$ due to $\mathbf{y} \in \mathcal{R}_{L,K}$, see (3.22).

Using $(1-\varphi)w_j - y_k \geq (1-\varphi)(w_j - y_k)$ for $L - 2K \leq k \leq L - K - 1$ and the identity

$$(1-\varphi)w_j - y_k = (w_j - y_k)\Big[1 - \frac{\varphi w_j}{w_j - y_k}\Big]$$

for any k, we have

$$\prod_{k\le L-K-1}((1-\varphi)w_j-y_k)^\beta \ge (1-\varphi)^{\beta K}\prod_{k\le L-K-1}(w_j-y_k)^\beta \prod_{n<L-2K}\Big[1-\frac{\varphi w_j}{w_j-y_n}\Big]^\beta, \qquad (6.28)$$

and a similar estimate holds for $k\ge L+K+1$. After multiplying these estimates for all $j=1,2,\ldots,K$, we thus have the bound

$$\frac{Z_\varphi}{Z_0}\ge\left[e^{-C\beta K\varphi}(1-\varphi)^{\beta K}\min_{|w|\le a}\left(\prod_{k<L-2K}\Big[1-\frac{\varphi w}{w-y_k}\Big]^\beta\prod_{k>L+2K}\Big[1-\frac{\varphi w}{y_k-w}\Big]^\beta\right)\right]^K. \qquad (6.29)$$

Recall that $\mathbf{y}\in\mathcal{R}_{L,K}$, i.e. we have the rigidity bound for $\mathbf{y}$ with accuracy $N^{-1}K^\xi\ll K/N\sim a$, see (3.7), i.e. y_k's are regularly spaced on scale a or larger. Combining this with $|w|\le a\le CK/N$, we have

$$\sum_{k\le L-2K}\frac{\varphi w}{w-y_k}\le C\varphi K\log N. \qquad (6.30)$$

Hence

$$\prod_{k<L-2K}\Big[1-\frac{\varphi w}{w-y_k}\Big]^\beta\ge 1-C\varphi K\log N, \qquad (6.31)$$

and similar bounds hold for the $k\ge L+2K$ factors. Thus for any $\varphi\le c$ we get

$$\frac{Z_\varphi}{Z_0}\ge 1-C\big(\beta K^2+K^2\log N\big)\varphi\ge 1-CK^2\varphi(\log N).$$

Now we choose $\varphi:=s/(aN)$ and recall $a\sim K/N$. Therefore the $\mu_{\mathbf{y}}$-probability of $x_{L+1}-y_L\ge a\varphi=s/N$ can be estimated by

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K}-y_{L-K-1}\ge s/N)\ge\frac{Z_\varphi}{Z_0}\ge 1-CKs(\log N).$$

for all $sK\log N$ sufficiently small. If $sK\log N$ is large, then (6.21) is automatically satisfied. This proves (6.21).

In order to prove (6.23), we now drop the assumption $\mathbf{y}\in\mathcal{R}_{L,K}$ and replace it with $|y_i|\le C$. Instead of (6.28), we now have

$$\prod_{k\le L-K-1}((1-\varphi)w_j-y_k)^\beta\ge(1-\varphi)^{\beta N}\prod_{k\le L-K-1}(w_j-y_k)^\beta, \qquad (6.32)$$

and a similar estimate holds for $k\ge L+K+1$. We thus have the bound

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K}-y_{L-K-1}\ge s/N)\ge\frac{Z_\varphi}{Z_0}\ge\left[e^{-C\beta K\varphi}(1-\varphi)^{\beta N}\right]^K\ge 1-C\varphi NK. \qquad (6.33)$$

With the choice $\varphi:=s/(|J_{\mathbf{y}}|N)$ this proves (6.23).

The proof of (6.22) and (6.24) for $\mu_{\mathbf{y},0}$ is very similar, just the $k=L-K-1$ factor is missing from (6.26) in case of $j=-K$. For $\mu_{\mathbf{y},1}$, two factors are missing. These modifications do not alter the basic estimates. This concludes the proof of Lemma 6.4. □

Proof of Lemma 6.3. Recalling the definition of μ_0 and setting $X := x_{L-K} - y_{L-K-1}$ for brevity, we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}[X \le s/N] = \frac{\mathbb{E}^{\mu_0}[1(X \le s/N)X^\beta]}{\mathbb{E}^{\mu_0}[X^\beta]}. \tag{6.34}$$

From (6.22) we have

$$\mathbb{E}^{\mu_0}[\mathbf{1}(X \le s/N)X^\beta] \le C(s/N)^\beta Ks \log N$$

and with the choice $s = cK^{-1}(\log N)^{-1}$ in (6.22) we also have

$$\mathbb{P}^{\mu_0}\left(X \ge \frac{c}{NK\log N}\right) \ge 1/2$$

with some positive constant c. This implies that

$$\mathbb{E}^{\mu_0}[X^\beta] \ge \frac{1}{2}\left(\frac{c}{NK\log N}\right)^\beta.$$

We have thus proved that

$$\mathbb{P}^{\mu_{\mathbf{y}}}[X \le s/N] \le C(s/N)^\beta Ks \log N\,(NK\log N)^\beta = C\,(Ks\log N)^{\beta+1}, \tag{6.35}$$

i.e. we obtained (6.17).

For the proof of (6.18), we similarly use

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{L-K+1} - y_{L-K-1} \le s/N] = \frac{\mathbb{E}^{\mu_1}[1(x_{L-K+1} - y_{L-K-1} \le s/N)W^\beta]}{\mathbb{E}^{\mu_1}[W^\beta]}. \tag{6.36}$$

From (6.22) we have

$$\mathbb{E}^{\mu_1}[\mathbf{1}(x_{L-K+1}-y_{L-K-1} \le s/N)W^\beta] \le (s/N)^{2\beta}\mathbb{P}^{\mu_1}[x_{L-K}-y_{L-K-1} \le s/N] \le C(s/N)^{2\beta}Ks\log N.$$

By the same inequality and with the choice $s = cK^{-1}(\log N)^{-1}$, we have

$$\mathbb{P}^{\mu_1}\left(W \ge \frac{c}{(NK\log N)^2}\right) \ge 1/2$$

with some positive constant c. This implies that

$$\mathbb{E}^{\mu_1}[W^\beta] \ge \frac{1}{2}\left(\frac{c}{(NK\log N)^2}\right)^\beta.$$

We have thus proved that

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{L-K+1} - y_{L-K-1} \le s/N] \le C(s/N)^{2\beta}Ks\log N\left((NK\log N)^2\right)^\beta = C\,(Ks\log N)^{2\beta+1}, \tag{6.37}$$

which proves (6.18). Finally, (6.19) and (6.20) can be proved using (6.23) and (6.24). This completes the proof of Lemma 6.3. □

Proof of Theorem 3.3. For a given i, define the index set

$$\widetilde{I} := [\![\max(i - K^\xi, L-K-1), \min(i + K^\xi, L+K+1)]\!]$$

to be the indices in a K^ξ neighborhood of i. We further condition the measure $\mu_{\mathbf{y}}$ on the points

$$z_j := x_j \qquad j \in \widetilde{I}^c := I_{L,K} \setminus \widetilde{I}$$

and we let $\mu_{\mathbf{y},\mathbf{z}}$ denote the conditional measure on the remaining x variables $\{x_j \; : \; j \in \widetilde{I}\}$. Setting $L' = i$, $K' = K^\xi$, from the rigidity estimate (3.16) we have $(\mathbf{y}, \mathbf{z}) \in \mathcal{R} = \mathcal{R}_{L',K'}(\xi^2\delta/2, \alpha)$ with a very high probability w.r.t. $\mu_{\mathbf{y}}$. We will now apply (6.17) to the measure $\mu_{\mathbf{y},\mathbf{z}}$ with a new $\delta' = \delta\xi$ and $K' = K^\xi$. This ensures that the condition $N^{\delta'} \le K'$ is satisfied and by the remark after (3.1), the change of δ affects only the threshold N_0. We obtain

$$\mathbb{P}^{\mu_{\mathbf{y},\mathbf{z}}}[x_i - x_{i+1} \le s/N] \le C\left(K^\xi s \log N\right)^{\beta+1} \tag{6.38}$$

with a high probability in $\mathbf{z}$ w.r.t. $\mu_{\mathbf{y}}$. The subexponential lower bound on s, assumed in part ii) of Theorem 3.3, allows us to include the probability of the complement of $\mathcal{R}$ in the estimate, we thus have proved (3.17). Similar argument but with (6.17) replaced by (6.18) yields (3.18).

To prove the weaker bounds (3.14), (3.15) for any $s > 0$, we may assume that $L - K \le i \le L$; $i > L$ is treated similarly. Since $\mathbf{y} \in \mathcal{R}_{L,K}$, we have $|J_{\mathbf{y}}| \ge cK/N$. We consider two cases, either $x_i - y_{L-K-1} \le c'K/N$ or $x_i - y_{L-K-1} \ge c'K/N$ with $c' < c/2$. In the first case, we condition on $x_{L-K}, \ldots, x_i$ and we apply (6.23) to the measure $\nu_1 = \mu_{\mathbf{y},x_{L-K},\ldots x_i}$. The configuration interval of this measure has length at least $cK/(2N)$, so we have

$$\mathbb{P}^{\nu_1}(x_{i+1} - x_i \le s/N) \le \frac{CKs}{cK/(2N)} \le CNs. \tag{6.39}$$

In the second case, $x_i - y_{L-K-1} \ge c'K/N$, we condition on $x_{i+1}, x_{i+2}, \ldots x_{L+K}$. The corresponding measure, denoted by $\nu_2 = \mu_{\mathbf{y},x_{i+1},\ldots x_{L+K}}$, has a configuration interval of length at least $c'K/N$. We can now have the estimate (6.39) for ν_2. Putting these two estimates together, we have proved (3.14). Finally (3.15) can be proved in a similar way. This completes the proof of Theorem 3.3.

□

6.3 Comparison of the local statistics of two local measures; proof of Theorem 3.1

In this section, we start to compare gap distributions of two local log-gases on the same configuration interval but with different external potential and boundary conditions. We will express the differences of gap distributions between two measures in terms of random walks in time dependent random environments. From now on, we use microscopic coordinates and we relabel the indices so that the coordinates of x_j are $j \in I = \{-K, \ldots, 0, 1, \ldots K\}$, i.e. we set $L = \widetilde{L} = 0$ in the earlier notation. This will have the effect that the labelling of the external points $\mathbf{y}$ will not run from 1 to N, but from some $L_- < 0$ to $L_+ > 0$ with $L_+ - L_- = N$. The important input is that the index set I of the internal points is macroscopically separated away from the edges, i.e. $|L_\pm| \ge \alpha N$.

The local equilibrium measures and their Hamiltonians will be denoted by the same symbols, $\mu_{\mathbf{y}}$ and $\mathcal{H}_{\mathbf{y}}$, as before, but with a slight abuse of notations we redefine them now to the microscopic scaling. Hence we have two measures $\mu_{\mathbf{y}} = e^{-\beta H_{\mathbf{y}}}/Z_{\mathbf{y}}$ and $\mu_{\widetilde{\mathbf{y}}} = e^{-\beta \widetilde{H}_{\widetilde{\mathbf{y}}}}/Z_{\widetilde{\mathbf{y}}}$, defined on the same configuration interval $J = J_{\mathbf{y}} = J_{\widetilde{\mathbf{y}}}$ with center $\bar{y}$, which, for simplicity, we assumed $\bar{y} = 0$. The local density at the center is $\varrho(0) > 0$. The Hamiltonian is given by

$$\mathcal{H}_{\mathbf{y}}(\mathbf{x}) := \sum_{i \in I} \frac{1}{2} V_{\mathbf{y}}(x_i) - \sum_{\substack{i,j \in I \\ i<j}} \log|x_j - x_i|$$

$$V_{\mathbf{y}}(x) := NV(x/N) - 2\sum_{j\notin I} \log|x-y_j|, \tag{6.40}$$

and $\widetilde{H}_{\tilde{\mathbf{y}}}$ is defined in a similar way with V in (6.40) replaced with another external potential $\widetilde{V}$. Recall also the assumption that $V'', \widetilde{V}'' \geq -C$ (2.10). We will need the rescaled version of the bounds (3.22), (3.23) and (3.24), i.e.

$$|J_{\mathbf{y}}| = \frac{\mathcal{K}}{\varrho(0)} + O(K^\xi), \tag{6.41}$$

$$V_{\mathbf{y}}'(x) = \varrho(0)\log\frac{d_+(x)}{d_-(x)} + O\Big(\frac{K^\xi}{d(x)}\Big), \qquad x\in J, \tag{6.42}$$

$$V_{\mathbf{y}}''(x) \geq \frac{\inf V''}{N} + \frac{c}{d(x)}, \qquad x\in J, \tag{6.43}$$

where

$$d(x) := \min\{|x-y_{-K-1}|, |x-y_{K+1}|\} \tag{6.44}$$

is the distance to the boundary and we redefined $d_\pm(x)$ as

$$d_-(x) := d(x) + \varrho(0)K^\xi, \qquad d_+(x) := \max\{|x-y_{-K-1}|, |x-y_{K+1}|\} + \varrho(0)K^\xi.$$

The rescaled version of Lemma 3.5 states that (6.41), (6.42) and (6.43) hold for any $\mathbf{y}\in\mathcal{R}_{L,K}(\xi\delta/2, \alpha/2)$, where the set $\mathcal{R}_{L,K}$, originally defined in (3.7), is expressed in microscopic coordinates.

We also rewrite (3.10) in the microscopic coordinate as

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j| + |\mathbb{E}^{\widetilde{\mu}_{\tilde{\mathbf{y}}}} x_j - \alpha_j| \leq CK^\xi, \tag{6.45}$$

where

$$\alpha_j := \frac{j}{\mathcal{K}+1}|J| \tag{6.46}$$

is the rescaled version of the definition given in (3.4), but we keep the same notation.

The Dirichlet form is also redefined; in microscopic coordinates it is now given by

$$D^{\mu_{\mathbf{y}}}(\sqrt{g}) = \sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{g}) = \sum_{i\in I}\int |\partial_i\sqrt{g}|^2 \mathrm{d}\mu_{\mathbf{y}}. \tag{6.47}$$

Due to the rescaling, the LSI from (5.16) now takes the form, for $\mathbf{y}\in\mathcal{R}_{L,K}$,

$$S(g\mu_{\mathbf{y}}|\mu_{\mathbf{y}}) \leq CKD^{\mu_{\mathbf{y}}}(\sqrt{g}). \tag{6.48}$$

Define the interpolating measures

$$\omega^r_{\mathbf{y},\tilde{\mathbf{y}}} = Z_r e^{-\beta r(\widetilde{V}_{\tilde{\mathbf{y}}}(\mathbf{x}) - V_{\mathbf{y}}(\mathbf{x}))}\mu_{\mathbf{y}}, \qquad r\in[0,1], \tag{6.49}$$

so that $\omega^1_{\mathbf{y},\tilde{\mathbf{y}}} = \widetilde{\mu}_{\tilde{\mathbf{y}}}$ and $\omega^0_{\mathbf{y},\tilde{\mathbf{y}}} = \mu_{\mathbf{y}}$ (Z_r is a normalization constant). This is again a local log-gas with Hamiltonian

$$\mathcal{H}^r_{\mathbf{y},\tilde{\mathbf{y}}}(\mathbf{x}) = \frac{1}{2}\sum_{i\in I} V^r_{\mathbf{y},\tilde{\mathbf{y}}}(x_i) - \frac{1}{N}\sum_{i<j}\log|x_i - x_j| \tag{6.50}$$

and external potential

$$
\begin{aligned}
V^r_{\mathbf{y},\widetilde{\mathbf{y}}}(x) &:= (1-r)V_{\mathbf{y}}(x) + r\widetilde{V}_{\widetilde{\mathbf{y}}}(x)\\
V_{\mathbf{y}}(x) &:= NV(x/N) - 2\sum_{j\notin I}\log(x-y_i),\\
\widetilde{V}_{\widetilde{\mathbf{y}}}(x) &:= N\widetilde{V}(x/N) - 2\sum_{j\notin I}\log(x-\widetilde{y}_i).
\end{aligned}
$$

The Dirichlet for D^ω w.r.t. the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is defined similarly to (6.47).

For any bounded smooth function O with compact support

$$
\partial_r \langle O(x_p - x_{p+1}, \dots x_p - x_{p+n})\rangle_{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} = \beta\langle h_0; O(x_p - x_{p+1}, \dots x_p - x_{p+n})\rangle_{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}}, \tag{6.51}
$$

where

$$
h_0 = h_0(\mathbf{x}) = \sum_{i\in I}(V_{\mathbf{y}}(x_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(x_i)) \tag{6.52}
$$

and $\langle f; g\rangle_\omega := \mathbb{E}^\omega fg - (\mathbb{E}^\omega f)(\mathbb{E}^\omega g)$ denotes the correlation. From now on, we will fix r. Our main result is the following estimate on the gap correlation function.

Theorem 6.5 *Consider two smooth potentials $V, \widetilde{V}$ with $V'', \widetilde{V}'' \ge -C$ and two boundary conditions, $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K}(\xi^2\delta/2, \alpha)$, with some sufficiently small ξ, such that $J = J_{\mathbf{y}} = J_{\widetilde{\mathbf{y}}}$. Assume that* (6.45) *holds for both boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}}$. Then, in particular, the rescaled version of the rigidity bound* (3.13) *and the level repulsion bounds* (3.17), (3.18) *hold for both $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ by Theorem 3.2 and Theorem 3.3.*

Fix $\xi^* > 0$. Then there exist $\varepsilon > 0$ and $C > 0$, depending on ξ^*, such that for any sufficiently small ξ, for any $0 \le r \le 1$ and for $|p| \le K^{1-\xi^*}$ we have

$$
|\langle h_0; O(x_p - x_{p+1}, \dots x_p - x_{p+n})\rangle_{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}}| \le K^{C\xi}K^{-\varepsilon} \tag{6.53}
$$

for any smooth function $O : \mathbb{R}^n \to \mathbb{R}$ with compact support provided that $K \ge K_0(\xi, \xi^*, n)$ is large enough.

Notice that this theorem is formulated in terms of K being the only large parameter; N disappeared. We also remark that the restriction $|p| \le K^{1-\xi^*}$ can be easily relaxed to $|p| \le K - K^{1-\xi^*}$ with an additional argument conditioning on set $\{x_i \,:\, i \in I \setminus \widetilde{I}\}$ to ensure that p is near the middle of the new index set $\widetilde{I}$. We will not need this more general form in this paper.

First we complete the proof of Theorem 3.1 assuming Theorem 6.5.

Proof of Theorem 3.1. The family of measures $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$, $0 \le r \le 1$, interpolate between $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$. So we can express the right hand side of (3.11), in the rescaled coordinates and with $L = \widetilde{L} = 0$ as

$$
\Big|[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}}]O(x_p - x_{p+1}, \dots x_p - x_{p+n})\Big| \le \int_0^1 \mathrm{d}r \frac{\mathrm{d}}{\mathrm{d}r}\mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} O(x_p - x_{p+1}, \dots x_p - x_{p+n}). \tag{6.54}
$$

Using (6.51) and (6.53) we obtain that this difference is bounded by $K^{C\xi}K^{-\varepsilon}$. Choosing ξ sufficiently small so that $K^{C\xi}K^{-\varepsilon} \le K^{-\varepsilon/2}$, we obtain (3.11) (with $\varepsilon/2$ instead of ε). This completes the proof of Theorem 3.1. $\square$

In the rest of the paper we will prove Theorem 6.5. The main difficulty is due to the fact that the correlation function of the points, $\langle x_i; x_j\rangle_\omega$, decays only logarithmically. In fact, for the GUE, Gustavsson proved that (Theorem 1.3 in [33])

$$\langle x_i; x_j\rangle_{GUE} \sim \log \frac{N}{[|i-j|+1]}, \tag{6.55}$$

and a similar formula is expected for ω. Therefore, it is very difficult to prove Theorem 6.5 based on this slow logarithmic decay. We notice that, however, the correlation function of the type

$$\langle g_1(x_i); g_2(x_j - x_{j+1})\rangle_\omega \tag{6.56}$$

decays much faster in $|i-j|$ due to that the second factor $g_2(x_j - x_{j+1})$ depends only on the difference. Correlations of the form $\langle g_1(x_i - x_{i+1}); g_2(x_j - x_{j+1})\rangle_\omega$ decay even faster. The fact that observables of differences of particles behave much nicer was a basic observation in our previous approach [21, 27, 28] of universality.

The measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is closely related to the measures $\mu_{\mathbf{y}}$ and $\mu_{\widetilde{\mathbf{y}}}$. Our first task in Section 6.4 is to show that both the rigidity and level repulsion estimates hold w.r.t. the measure ω. Then we will rewrite the correlation functions in terms of a random walk representation in Proposition 7.1. The decay of correlation functions will be translated into a partial regularity property of the corresponding parabolic equation, whose proof will be the main content of Section 8. Section 7 consists of various cutoff estimates to remove the singularity of the diffusion coefficients in the random walk representations. We emphasize that these cutoffs are critical at $\beta = 1$; we do not know if our argument can be extended to $\beta < 1$.

6.4 Rigidity and level repulsion of the interpolating measure $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$

In this section we establish rigidity and level repulsion results for the interpolating measure $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$, similar to the ones established for $\mu_{\mathbf{y}}$ in Section 6 and stated in Theorems 3.2 and 3.3.

Lemma 6.6 *Let L and K satisfy* (3.1) *and $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$. With the notation $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ there exist constants C, θ_3, C_2 and C_3 such that the following estimates hold:*

i) [Rigidity bound]

$$\mathbb{P}^\omega\big(|x_i - \alpha_i| \ge CK^{C_2\xi^2}\big) \le Ce^{-K^{\theta_3}}, \quad i \in I. \tag{6.57}$$

ii) [Weak form of level repulsion] For any $s > 0$ we have

$$\mathbb{P}^\omega\big(x_{i+1} - x_i \le s\big) \le C\,(Ns)^{\beta+1}, \quad i \in [\![L-K-1, L+K]\!], \quad s > 0, \tag{6.58}$$

$$\mathbb{P}^\omega\big(x_{i+2} - x_i \le s\big) \le C\,(Ns)^{2\beta+1}, \quad i \in [\![L-K-1, L+K-1]\!], \quad s > 0, \tag{6.59}$$

iii) [Strong form of level repulsion] With some small $\theta > 0$, for any $s \ge \exp(-K^\theta)$ we have

$$\mathbb{P}^\omega\big(x_{i+1} - x_i \le s\big) \le C\,\big(K^{C_3\xi}s\big)^{\beta+1}, \quad i \in [\![L-K-1, L+K]\!], \tag{6.60}$$

$$\mathbb{P}^\omega\big(x_{i+2} - x_i \le s\big) \le C\,\big(K^{C_3\xi}s\big)^{2\beta+1}, \quad i \in [\![L-K-1, L+K-1]\!], \tag{6.61}$$

iv) [Logarithmic Sobolev inequality]

$$S(g\omega|\omega) \le CKD^\omega(\sqrt{g}). \tag{6.62}$$

Note that in (6.57) we state only the weaker form of the rigidity bound, similar to (3.16). It is possible to prove the strong form of rigidity with Gaussian tail (3.13) for ω, but we will not need it in this paper.

The level repulsion bounds will mostly be used in the following estimates which trivially follow from (6.58)–(6.61):

Corollary 6.7 *Under the assumptions of Lemma 6.6, for any $p < \beta + 1$ we have*

$$\mathbb{E}^{\omega} \frac{1}{|x_i - x_{i+1}|^p} \leq C_p K^{C_3 \xi}, \qquad i \in [\![L-K-1, L+K]\!], \tag{6.63}$$

and for any $p < 2\beta + 1$

$$\mathbb{E}^{\omega} \frac{1}{|x_i - x_{i+2}|^p} \leq C_p K^{C_3 \xi}, \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{6.64}$$

□

The key to translate the rigidity estimate of the measures $\mu_{\mathbf{y}}$ and $\mu_{\widetilde{\mathbf{y}}}$ to the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is to show that the analogue of (6.45) holds for ω.

Lemma 6.8 *Let L and K satisfy* (3.1) *and $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in* (3.6) *and assume that* (3.10) *is satisfied. Let $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ be the measure defined in* (6.49)*. Recall that α_k denote the equidistant points in J, see* (6.46)*. Then there exists a constant C, independent of ξ, such that*

$$\mathbb{E}^{\omega^r_{\mathbf{y},\tilde{\mathbf{y}}}} |x_j - \alpha_j| \leq C K^{C\xi}. \tag{6.65}$$

Proof of Lemma 6.8. We first prove the following estimate on the entropy.

Lemma 6.9 *Suppose μ_1 is a probability measure and $\omega = Z^{-1} e^g d\mu_1$ for some function g and normalization Z. Then we can bound the entropy by*

$$S := S(\omega|\mu_1) = \mathbb{E}^{\omega} g - \log \mathbb{E}^{\mu_1} e^g \leq \mathbb{E}^{\omega} g - \mathbb{E}^{\mu_1} g. \tag{6.66}$$

Consider two probability measures $\mathrm{d}\mu_i = Z_i^{-1} e^{-H_i} \mathrm{d}\mathbf{x}$, $i = 1, 2$. Denote by g the function

$$g = r(H_1 - H_2), \quad 0 < r < 1 \tag{6.67}$$

and set $\omega = Z^{-1} e^g d\mu_1$ as above. Then we can bound the entropy by

$$\min(S(\omega|\mu_1), S(\omega|\mu_2)) \leq \Big[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1} \Big] (H_1 - H_2). \tag{6.68}$$

Proof. The first inequality is a trivial consequence of the Jensen inequality

$$S = \mathbb{E}^{\omega} g - \log \mathbb{E}^{\mu_1} e^g \leq \mathbb{E}^{\omega} g - \mathbb{E}^{\mu_1} g.$$

The entropy inequality yields that

$$\mathbb{E}^{\omega} g \leq r \log \mathbb{E}^{\mu_1} e^{g/r} + rS. \tag{6.69}$$

By the definition of g, we have

$$\log \mathbb{E}^{\mu_1} e^{g/r} = -\log \int e^{-g/r} \mathrm{d}\mu_2 \leq \mathbb{E}^{\mu_2} g/r.$$

Using this inequality and (6.69) in (6.66), we have proved

$$S \le \frac{r}{1-r}\Big[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}\Big](H_1 - H_2). \tag{6.70}$$

We can assume that $r \le 1/2 \le 1-r$ since otherwise we can switch the roles of H_1 and H_2. Hence (6.68) holds and this concludes the proof of Lemma 6.9. □

We now apply this lemma with $\mu_2 = \widetilde{\mu}_{\widetilde{\mathbf{y}}}$ and $\mu_1 = \mu_{\mathbf{y}}$ to prove that

$$\min[S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\mu_{\mathbf{y}}), S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\widetilde{\mu}_{\widetilde{\mathbf{y}}})] \le K^{C\xi} \tag{6.71}$$

To see this, by definition of g and the rigidity estimate (3.13), we have

$$\begin{aligned}
\mathbb{E}^{\mu_2} g - \mathbb{E}^{\mu_1} g &= \frac{r}{2}\Big[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}\Big]\sum_{i\in I}\Big[V_{\mathbf{y}}(x_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(x_i)\Big] \\
&= \frac{r}{2}\Big[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}\Big]\sum_{i\in I}\int_0^1 \mathrm{d}s\Big[V'_{\mathbf{y}}(s\alpha_i + (1-s)x_i) - \widetilde{V}'_{\widetilde{\mathbf{y}}}(s\alpha_i + (1-s)x_i)\Big](x_i - \alpha_i) \\
&= \Big[\mathbb{E}^{\mu_2} + \mathbb{E}^{\mu_1}\Big] O\Big(\sum_{i\in I}\sup_{s\in[0,1]}\frac{K^\xi}{d(s\alpha_i + (1-s)x_i)}|x_i - \alpha_i|\Big) \le K^{C\xi}.
\end{aligned} \tag{6.72}$$

In the first step we used that the leading term $V_{\mathbf{y}}(\alpha_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(\alpha_i)$ in the Taylor expansion is deterministic, so it vanishes after taking the difference of two expectations. In the last step we used that with a very high μ_1- or μ_2-probability $d(s\alpha_i + (1-s)x_i) \sim d(\alpha_i)$ are equidistant up to an additive error K^ξ if i is away from the boundary, i.e., $-K + K^{C\xi} \le i \le K - K^{C\xi}$, see (3.13). For indices near the boundary, say $-K \le i \le -K + K^{C\xi}$, we used $d(s\alpha_i + (1-s)x_i) \ge c\min\{1, d(x_{-K})\}$. Noticing that $d(x_{-K}) = x_{-K} - y_{-K-1}$, the level repulsion bound (3.17) (complemented with the weaker bound (6.17) that is valid for all $s > 0$) guarantees that the short distance singularity $[d(x_{-K})]^{-1}$ has an $\mathbb{E}^{\mu_{1,2}}$ expectation that is bounded by $CK^{C\xi}$.

We now assume that (6.71) holds with the choice of $S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\mu_{\mathbf{y}})$ for simplicity of notation. By the entropy inequality, we have

$$\mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}}|x_i - \alpha_i| \le \log \mathbb{E}^{\mu_{\mathbf{y}}} e^{|x_i - \alpha_i|} + K^{C\xi}. \tag{6.73}$$

From the Gaussian tail of the rigidity estimate (3.13), we have

$$\log \mathbb{E}^{\mu_{\mathbf{y}}} e^{|x_i - \alpha_i|} \le K^{C\xi}. \tag{6.74}$$

Using this bound in (6.73) we have proved (6.65) and this concludes the proof of Lemma 6.8. □

Proof of Lemma 6.6. Given (6.65), the proof of (6.57) follows the argument in the proof of Theorem 3.2, applying it to ξ^2 instead of ξ. Once the rigidity bound (6.57) is proved, we can follow the proof of Theorem 3.3 to obtain all four level repulsion estimates, (6.58)–(6.61), analogously to the proofs of (3.14), (3.15), (3.17) and (3.18), respectively. The $\log N$ factor can be incorporated into $K^{C_3\xi}$.

Finally, to prove (6.62), let $\mathcal{L}^\omega$ be the reversible generator given by the Dirichlet form

$$-\int f\mathcal{L}^\omega f \mathrm{d}\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}} = \sum_{|j|\le K}\int (\partial_j f)^2 \mathrm{d}\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}. \tag{6.75}$$

Thus for the Hamiltonian $\mathcal{H} = \mathcal{H}^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ of the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ (see (6.50)), we have

$$\Big\langle \mathbf{v}, \nabla^2\mathcal{H}(\mathbf{x})\mathbf{v}\Big\rangle = \frac{1}{2}\sum_i \Big[(1-r)V''_{\mathbf{y}}(x_i) + r\widetilde{V}''_{\widetilde{\mathbf{y}}}(x_i)\Big]v_i^2 + \sum_{i<j}\frac{(v_i-v_j)^2}{(x_i-x_j)^2} \geq \frac{c}{K}\sum_i v_i^2, \tag{6.76}$$

by using (6.43) and $d(x) \leq CK$ for good boundary conditions. Thus LSI takes the form

$$S(g\omega|\omega) \leq CKD^\omega(\sqrt{g}). \tag{6.77}$$

This completes the proof of Lemma 6.6. □

The dynamics given by the generator $\mathcal{L}^\omega$ with respect to the interpolating measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ can also be characterized by the following SDE

$$\mathrm{d}x_i = \mathrm{d}B_i + \beta\Big[-\frac{1}{2}(V^r_{\mathbf{y},\widetilde{\mathbf{y}}})'(x_i) + \frac{1}{2}\sum_{j\neq i}\frac{1}{(x_i-x_j)}\Big]\mathrm{d}t, \tag{6.78}$$

where $(B_{-K}, B_{-K+1}, \dots, B_K)$ is a family of independent standard Brownian motions. With a slight abuse of notations, when we talk about the process, we will use $\mathbb{P}^\omega$ and $\mathbb{E}^\omega$ to denote the probability and expectation w.r.t. this dynamics with initial data ω, i.e., in equilibrium. This dynamical point of view gives rise to a representation for the correlation function (6.53) in terms random walks in random environment.

Starting from Section 7 we will focus on proving Theorem 6.5. The proof is based on dynamical idea and it will be completed in Section 7.5.

7 Local statistics of the interpolating measures: Proof of Theorem 6.5

7.1 Random Walk Representation

In this section we derive a random walk representation for the gap correlation function on the left hand side of (6.53). For brevity of the formulas, we consider only the single gap case, $n = 1$; the general case is a straightforward extension. The gap index p is fixed, $-K \leq p \leq K-1$.

Suppose $h(t) = h(t,\mathbf{x})$ is the solution of the equation $\partial_t h = \mathcal{L}^\omega h$ where $\mathcal{L} = \mathcal{L}^\omega$ is the unique reversible generator with Dirichlet form $D^\omega(g)$ w.r.t. the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$. The initial condition is chosen to be h_0 given in (6.52). Introduce the notation

$$\mathbf{v}(t,\mathbf{x}) = \nabla_{\mathbf{x}}h(t,\mathbf{x}), \quad \text{i.e.} \quad v_j(t,\mathbf{x}) := \partial_{x_j}h(t,\mathbf{x}), \tag{7.1}$$

then

$$v_j(0,\mathbf{x}) = \partial_{x_j}h_0(\mathbf{x}) = (V_{\mathbf{y}})'(x_j) - (\widetilde{V}_{\widetilde{\mathbf{y}}})'(x_j), \quad j \in I. \tag{7.2}$$

By integrating the time derivative of $\langle h(t,\mathbf{x}); O(x_p - x_{p+1})\rangle_\omega$ and using the equation $h(t,\mathbf{x})$ satisfies, we have

$$\langle h_0(\mathbf{x}); O(x_p - x_{p+1})\rangle_\omega = \int_0^\infty \mathrm{d}\sigma \int O'(x_p - x_{p+1})[v_p(\sigma,\mathbf{x}) - v_{p+1}(\sigma,\mathbf{x})]\mathrm{d}\omega(\mathbf{x}). \tag{7.3}$$

Now we derive a random walk representation for (7.3). Taking the gradient of the equation $\partial_t h = \mathcal{L}h$, a direct computation of the commutator $[\nabla, \mathcal{L}]$ yields that

$$\partial_t \mathbf{v}(t, \mathbf{x}) = \mathcal{L}\mathbf{v}(t, \mathbf{x}) - \widetilde{\mathcal{A}}(\mathbf{x})\mathbf{v}(t, \mathbf{x}), \tag{7.4}$$

with initial condition $\mathbf{v}_0(\mathbf{x}) = \mathbf{v}(0, \mathbf{x}) = \nabla h_0(\mathbf{x})$, see (7.2). Here $\widetilde{\mathcal{A}} := \widetilde{\mathcal{B}} + \widetilde{\mathcal{W}}$ and $\widetilde{\mathcal{B}}$ and $\widetilde{\mathcal{W}}$ are the following $\mathbf{x}$-dependent operators acting on vectors $\mathbf{v} \in \mathbb{R}^{\mathcal{K}}$ with components v_i, $i \in I$:

$$[\widetilde{\mathcal{B}}(\mathbf{x})\mathbf{v}]_j = -\sum_k \widetilde{B}_{jk}(\mathbf{x})(v_k - v_j), \quad \widetilde{B}_{jk}(\mathbf{x}) = \frac{1}{(x_j - x_k)^2} \geq 0 \tag{7.5}$$

and

$$[\widetilde{\mathcal{W}}(\mathbf{x})\mathbf{v}]_j := \widetilde{W}_j(\mathbf{x})v_j \tag{7.6}$$

with the function

$$\widetilde{W}_j(\mathbf{x}) := \frac{1}{2}\Bigg\{ \sum_{|k| \geq K+1} \Big[\frac{1-r}{(x_j - y_k)^2} + \frac{r}{(x_j - \widetilde{y}_k)^2}\Big] + \frac{1-r}{N}V''\big(\frac{x_j}{N}\big) + \frac{r}{N}\widetilde{V}''\big(\frac{x_j}{N}\big)\Bigg\}. \tag{7.7}$$

Here $r \in [0,1]$ is a fixed parameter which we will omit from the notation of $\widetilde{\mathcal{W}}$.

Recall that $\mathbf{x}(t)$ is the stochastic process (6.78) with generator $\mathcal{L} = \mathcal{L}^\omega$ and let $\mathbb{E}_{\mathbf{x}}$ denote the expectation for this process with initial point $\mathbf{x}(0) = \mathbf{x}$. The expectation with respect to the process starting from equilibrium is $\mathbb{E}^\omega[\cdot] = \int \mathbb{E}_{\mathbf{x}}[\cdot]\omega(\mathrm{d}\mathbf{x})$. The solution to (7.4) can be represented by the Feynman-Kac formula:

$$\mathbf{v}(\sigma, \mathbf{x}) = \mathbb{E}_{\mathbf{x}}\, \widetilde{\mathrm{Exp}}\Big(- \int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\mathrm{d}s\Big)\mathbf{v}_0(\mathbf{x}(\sigma)), \tag{7.8}$$

where

$$\widetilde{\mathrm{Exp}}\Big(- \int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\mathrm{d}s\Big) := 1 - \int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s_1))\mathrm{d}s_1 + \int_{0 \leq s_1 < s_2 \leq \sigma} \widetilde{\mathcal{A}}(\mathbf{x}(s_1))\widetilde{\mathcal{A}}(\mathbf{x}(s_2))\mathrm{d}s_1\mathrm{d}s_2 + \dots \tag{7.9}$$

is the time-ordered exponential. To prove that (7.8) indeed satisfies (7.4), we notice from the definition (7.9) that

$$\begin{aligned} \mathbf{v}(\sigma, \mathbf{x}) &= \mathbb{E}_{\mathbf{x}}\, \widetilde{\mathrm{Exp}}\Big(- \int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\mathrm{d}s\Big)\mathbf{v}_0(\mathbf{x}(\sigma)) \\ &= \mathbb{E}_{\mathbf{x}}\mathbf{v}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}}\widetilde{\mathcal{A}}(\mathbf{x}(s_1))\mathbb{E}_{\mathbf{x}(s_1)}\widetilde{\mathrm{Exp}}\Big(\int_{s_1}^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\mathrm{d}s\Big)\mathbf{v}_0(\mathbf{x}(\sigma))\mathrm{d}s_1. \end{aligned} \tag{7.10}$$

Using that the process is stationary in time, we have

$$\begin{aligned} \mathbf{v}(\sigma, \mathbf{x}) &= \mathbb{E}_{\mathbf{x}}\mathbf{v}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}}\widetilde{\mathcal{A}}(\mathbf{x}(s_1))\mathbf{v}(\sigma - s_1, \mathbf{x}(s_1))\mathrm{d}s_1 \\ &= \mathbb{E}_{\mathbf{x}}\mathbf{v}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}}\widetilde{\mathcal{A}}(\mathbf{x}(\sigma - s_1))\mathbf{v}(s_1, \mathbf{x}(\sigma - s_1))\mathrm{d}s_1 \\ &= e^{\sigma\mathcal{L}}\mathbf{v}_0(\mathbf{x}) - \int_0^\sigma [e^{(\sigma - s_1)\mathcal{L}}\widetilde{\mathcal{A}}(\cdot)\mathbf{v}(s_1, \cdot)](\mathbf{x})\mathrm{d}s_1. \end{aligned}$$

Differentiating this equation in σ we obtain that $\mathbf{v}$ defined in (7.8) indeed satisfies (7.4).

For any path $\mathbf{x}(\cdot)$ starting from $\mathbf{x}$ fixed at time zero, $\mathbf{x}(0) = \mathbf{x}$, and for any vector $\mathbf{w}_0$, denote the vector

$$\widetilde{\mathbf{w}}(s; \mathbf{x}(\cdot), \sigma) := \widetilde{\operatorname{Exp}}\Big(- \int_s^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(b)) \mathrm{d}b \Big) \mathbf{w}_0. \tag{7.11}$$

Then we have

$$\begin{aligned} \partial_s \widetilde{\mathbf{w}}(s; \mathbf{x}(\cdot), \sigma) &= \partial_s \Big[1 - \int_s^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s_1)) \mathrm{d}s_1 + \int_{s \le s_1 < s_2 \le \sigma} \widetilde{\mathcal{A}}(\mathbf{x}(s_1)) \widetilde{\mathcal{A}}(\mathbf{x}(s_2)) \mathrm{d}s_1 \mathrm{d}s_2 - \dots \Big] \mathbf{w}_0 \\ &= \widetilde{\mathcal{A}}(\mathbf{x}(s)) \widetilde{\mathbf{w}}(s; \mathbf{x}(\cdot), \sigma). \end{aligned}$$

This is the backward equation, starting from $\mathbf{w}(s = \sigma; \mathbf{x}(\cdot), \sigma) = \mathbf{w}_0$ and evolving backwards in time from $s = \sigma$ to $s = 0$, for a random walk given by the random environment induced by the path $\mathbf{x}(\cdot)$. The equation in a slightly different setting already appeared in Proposition 2.2 of [16] (see also Proposition 3.1 in [32]), which was a probabilistic formulation of the idea of Helffer and Sjöstrand [35] and Naddaf and Spencer [41]. For our purpose, we will convert the backward equation into the forward equation in order to be consistent with the convention in the partial differential equation literature [10]. Thus replacing s with $\sigma - s$ in (7.11) we define

$$\mathbf{w}(s; \mathbf{x}(\cdot), \sigma) := \widetilde{\mathbf{w}}(\sigma - s; \mathbf{x}(\cdot), \sigma) = \widetilde{\operatorname{Exp}}\Big(- \int_{\sigma - s}^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(b)) \mathrm{d}b \Big) \mathbf{w}_0. \tag{7.12}$$

We have then the evolution equation

$$\partial_s \mathbf{w}(s; \mathbf{x}(\cdot), \sigma) = -\widetilde{\mathcal{A}}(\mathbf{x}(\sigma - s)) \mathbf{w}(s; \mathbf{x}(\cdot), \sigma), \quad \mathbf{w}(0; \mathbf{x}(\cdot), \sigma) = \mathbf{w}_0. \tag{7.13}$$

With these notations, and choosing $\mathbf{w}_0 = \mathbf{v}_0(\mathbf{x}(\sigma))$ in the initial value problem (7.13), we have from (7.8) that

$$\mathbf{v}(\sigma, \mathbf{x}) = \mathbb{E}_{\mathbf{x}} \mathbf{w}(\sigma; \mathbf{x}(\cdot), \sigma). \tag{7.14}$$

In other words, for any path starting from $\mathbf{x}$ and ending at $\mathbf{x}(\sigma)$, we solve (7.13), which is $\mathcal{K}$-dimensional system of ODE. Then taking expectation yields the solution $\mathbf{v}$ to (7.4).

We now slightly rewrite the equation (7.13) into a more convenient form. Fix $0 < \sigma \le C_1 K \log K$ and fix a path $\{\mathbf{x}(s) \; : \; s \in [0, \sigma]\}$. Define the following operators on $\mathbb{R}^{\mathcal{K}}$

$$\mathcal{A}(s) := \widetilde{\mathcal{A}}(\mathbf{x}(\sigma - s)), \quad \mathcal{B}(s) := \widetilde{\mathcal{B}}(\mathbf{x}(\sigma - s)), \quad \mathcal{W}(s) := \widetilde{\mathcal{W}}(\mathbf{x}(\sigma - s)), \tag{7.15}$$

where $\mathcal{W}$ is a multiplication operator with the j-th diagonal $W_j(s) = \widetilde{W}_j(x_j(\sigma - s))$ depending only the j-th component of the process $\mathbf{x}(s)$. Clearly $\mathcal{A}(s) = \mathcal{B}(s) + \mathcal{W}(s)$. We also define the associated (time dependent) quadratic forms which we denote by the corresponding lower case letters, in particular

$$\mathfrak{b}(s)[\mathbf{u}, \mathbf{v}] = \sum_{i \in I} u_i [\mathcal{B}(s) \mathbf{v}]_i = \frac{1}{2} \sum_{k, j \in I} B_{jk}(s)(u_k - u_j)(v_k - v_j)$$

$$\mathfrak{w}(s)[\mathbf{u}, \mathbf{v}] = \sum_{i \in I} u_i [\mathcal{W}(s) \mathbf{v}]_i = \sum_i u_i W_i(s) v_i$$

and

$$\mathfrak{a}(s)[\mathbf{u}, \mathbf{v}] = \mathfrak{b}(s)[\mathbf{u}, \mathbf{v}] + \mathfrak{w}(s)[\mathbf{u}, \mathbf{v}]. \tag{7.16}$$

With this notation, (7.13) becomes

$$\partial_s \mathbf{w}(s;\mathbf{x}(\cdot),\sigma) = -\mathcal{A}(s)\mathbf{w}(s;\mathbf{x}(\cdot),\sigma), \quad \mathbf{w}(0;\mathbf{x}(\cdot),\sigma) = \nabla h_0(\mathbf{x}(\sigma)). \tag{7.17}$$

Combining (7.14) and (7.17) with (7.3), we have proved the following proposition. Notice that in the above derivation we never used the explicit form of the initial condition h_0 given in (6.52), so we can formulate the result for general h_0.

Proposition 7.1 *For any smooth function $h_0 : J^{\mathcal{K}} \to \mathbb{R}$, for any $p \in I$, $-K \le p \le K-1$, we have*

$$\langle h_0; O(x_p - x_{p+1})\rangle_\omega = \int_0^\infty \mathrm{d}\sigma \int O'(x_p - x_{p+1})\mathbb{E}_{\mathbf{x}}[w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)]\omega(\mathrm{d}\mathbf{x}) \tag{7.18}$$

Here for any $\sigma > 0$ and for any path $\{\mathbf{x}(s) \in J^{\mathcal{K}} \ : \ s \in [0,\sigma]\}$ we let $\mathbf{w}$ be the solution of (7.17) with initial data $\mathbf{w}(0;\mathbf{x}(\cdot),\sigma) := \nabla h_0(\mathbf{x}(\sigma))$.

Now we apply this proposition to our case:

Corollary 7.2 *Let h_0 be given by (6.52) and assume that $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K}(\xi^2\delta/2,\alpha)$. Then for any $p \in I$, $-K \le p \le K-1$, we have*

$$\begin{aligned} &\langle h_0; O(x_p - x_{p+1})\rangle_\omega \\ &\quad = \int_0^{C_1 K \log K} \mathrm{d}\sigma \int O'(x_p - x_{p+1})\mathbb{E}_{\mathbf{x}}[w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)]\omega(\mathrm{d}\mathbf{x}) + O(K^{-2}). \end{aligned} \tag{7.19}$$

Proof. Since $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K}(\xi^2\delta/2,\alpha)$, we easily obtain that $W_j(s) \ge cK^{-1}$, which means that the L^1-norm of the solution to (7.17) decays at a rate of order K. Therefore the integral in (7.18) from $C_1 K \log K$ to infinity can be bounded by

$$C\int_{C_1 K \log K}^\infty e^{-c\sigma/K}\mathbb{E}^\omega |v_p(\sigma,\cdot)|\mathrm{d}\sigma \le CK^{(C_3+1)\xi}K^{-C_1} \le K^{-2}$$

if the constant C_1 is chosen sufficiently large. Here we used that

$$\begin{aligned} \mathbb{E}^\omega |v_p(\sigma,\cdot)| = \mathbb{E}^\omega |v_p(0,\cdot)| &\le \mathbb{E}^\omega\big[|(V_{\mathbf{y}})'(x_p)| + |(V_{\widetilde{\mathbf{y}}})'(x_p)|\big] \\ &\le \sum_{j \notin I}\mathbb{E}^\omega\Big[\frac{1}{|y_j - x_p|} + \frac{1}{|\widetilde{y}_j - x_p|}\Big] \le CK^{(C_3+1)\xi} \end{aligned} \tag{7.20}$$

from (6.63) and using that $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K} = \mathcal{R}_{L=0,K}(\xi^2\delta/2,\alpha)$ are regular on scale $K^{\xi^2} \le K^\xi$, so the summation is effectively restricted to K^ξ terms. □

The representation (7.18) and (7.19) expresses the correlation function in terms of the discrete spatial derivative of the solution to (7.17). To estimate $w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)$ in (7.19), we will now study the Hölder continuity of the solution $\mathbf{w}(s,\mathbf{x}(\cdot);\sigma)$ to (7.17) at time $s = \sigma$ and at the spatial point p. We will do it for each fixed path $\mathbf{x}(\cdot)$, with the exception of a set of "bad" paths that will have a small probability.

Notice that if all points x_i were approximately regularly spaced in the interval J, then the operator $\mathcal{B}$ has a kernel $B_{ij} \sim (i-j)^{-2}$, i.e. it is essentially a discrete version of the operator $|p| = \sqrt{-\Delta}$. Hölder continuity will thus be the consequence of the De Giorgi-Nash-Moser bound

for the parabolic equation (7.17). However, we need to control the coefficients in this equation, which depend on the random walk $\mathbf{x}(\cdot)$.

For the De Giorgi-Nash-Moser theory we need both upper and lower bounds on the kernel B_{ij}. The rigidity bound (6.57) guarantees a lower bound on B_{ij}, up to a factor $K^{-C_2\xi^2} \geq K^{-\xi}$. The level repulsion estimate implies certain upper bounds on B_{ij}. In the next section we define the good set of paths that satisfy both requirements.

7.2 Sets of good paths

From now on we assume the conditions of Theorem 6.5. In particular we are given some $\xi > 0$ and we assume that the boundary conditions satisfy $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K} = \mathcal{R}_{L=0,K}(\xi^2\delta/2, \alpha)$ and (6.45) with this ξ. We define the following "good sets":

$$\mathcal{G} := \Big\{ \sup_{0 \leq s \leq C_1 K \log K} \sup_{|j| \leq K} |x_j(s) - \alpha_j| \leq K^{\xi'} \Big\}, \tag{7.21}$$

where

$$\xi' := (C_2 + 1)\xi^2,$$

with C_2 being the constant in (6.57) and α_j is given by (6.46). For any $Z \in I$ and $\sigma \in [0, C_1 K \log K]$ we also define the event

$$\begin{aligned} \mathcal{Q}_{\sigma,Z} := \Big\{ \sup_{0 \leq s \leq \sigma} \sup_{1 \leq M \leq K} \frac{1}{1+s} \int_0^s \mathrm{d}a \, \frac{1}{M} \\ \sum_{i=\max\{-M+Z,-K-1\}}^{\min\{M+Z,K\}} \frac{1}{\big|x_i(\sigma - a) - x_{i+1}(\sigma - a)\big|^2} \leq K^\rho \Big\}, \end{aligned} \tag{7.22}$$

where $\rho > 0$ will be specified later and, by convention, $x_i(\sigma - a) = y_i$ whenever $|i| > K$. Define a new good set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ by

$$\widetilde{\mathcal{Q}}_{\sigma,Z} = \mathcal{Q}_{\sigma,Z} \cap \mathcal{Q}_{\sigma,-K} \cap \mathcal{Q}_{\sigma,K}, \tag{7.23}$$

i.e., this is the event that the gaps between particles near Z and near the boundary are not too small in an appropriate average sense.

Lemma 7.3 *There exists a positive constant θ, depending on $\xi' = (C_2 + 1)\xi^2$, such that*

$$\mathbb{P}^\omega(\mathcal{G}^c) \leq Ce^{-K^\theta}. \tag{7.24}$$

Moreover, there is a constant C_4, depending on the constant C_2 in (6.57) *and on C_3 in* (6.60), (6.61) *such that for any ξ and ρ small enough, we have*

$$\mathbb{P}^\omega(\widetilde{\mathcal{Q}}^c_{\sigma,Z}) \leq CK^{C_4\xi - \rho} \tag{7.25}$$

for each fixed $Z \in I$ and fixed $\sigma \in [0, C_1 K \log K]$.

Proof. From the stochastic differential equation of the dynamics (6.78) we have

$$|x_i(t) - x_i(s)| \leq C|t-s| + \int_s^t \Big[\sum_{\substack{j \in I \\ j \neq i}} \frac{1}{|x_j(a) - x_i(a)|} + \sum_{j \in I^c} \frac{1}{|y_j - x_i(a)|} \Big] \mathrm{d}a + |B_i(t) - B_i(s)|. \tag{7.26}$$

Using (6.63) and that $\mathbf{x}(\cdot)$ is invariant under ω, we have the bound

$$\mathbb{E}^\omega\Big[\int_s^t \sum_{j\neq i}\frac{1}{|x_j(a)-x_i(a)|}\Big]^{3/2} \le CK^3|t-s|^{3/2}\mathbb{E}^\omega\frac{1}{|x_j-x_i|^{3/2}} \le CK^{3+C_3\xi}|t-s|^{3/2}. \qquad (7.27)$$

This implies for any fixed $s<t\le C_1K\log K$ and for any $R>0$ that

$$\mathbb{P}^\omega\left[\int_s^t \sum_{j\neq i}\frac{1}{|x_j(a)-x_i(a)|}\ge R\right] \le CK^{3+C_3\xi}|t-s|^{3/2}R^{-3/2}. \qquad (7.28)$$

A similar bound holds for the second summation in (7.26); the summation over large j can be performed by using that $\mathbf{y}$ is regular, $\mathbf{y}\in\mathcal{R}_{L=0,K}$.

Set a parameter $q\le cR$ and choose a discrete set of increasing times $\{s_k\ :\ k=0,1,2\ldots,K^2/q\}$ such that

$$0=s_0<s_1\le s_2\le\ldots\le C_1K\log K, \qquad \text{and} \quad |s_k-s_{k+1}|\le q. \qquad (7.29)$$

From standard large deviation bounds on the Brownian motion increment $B_i(t)-B_i(s)$ and from (7.26), we have the stochastic continuity estimate

$$\mathbb{P}^\omega\left(\sup_{s,t\in[s_k,s_{k+1}],|i|\le K}|x_i(s)-x_i(t)|\ge R\right)\le Ke^{-CR^2/q}+CK^4q^{3/2}R^{-3/2}$$

for any fixed k. Taking sup over k, we have

$$\mathbb{P}^\omega\left(\sup_{0\le s,t\le C_1K\log K,|t-s|\le q,|i|\le K}|x_i(s)-x_i(t)|\ge R\right)\le K^3q^{-1}e^{-CR^2/q}+CK^6q^{1/2}R^{-3/2} \qquad (7.30)$$

for any positive q and R with $q\le cR$.

From the rigidity bound (6.57) we know that for some $\theta_3>0$ and for any fixed k we have,

$$\mathbb{P}^\omega\Big\{|x_j(s_k)-\alpha_j|\ge CK^{C_2\xi^2}\Big\}\le Ce^{-K^{\theta_3}}, \quad j\in I. \qquad (7.31)$$

Choosing $R=K^{\xi'}/2$ and $q=\exp(-K^{\theta_3/2})$, and using that $CK^{C_2\xi^2}\le K^{\xi'}/2$ with the choice of ξ', we have

$$\mathbb{P}^\omega(\mathcal{G}^c)\le Ce^{-K^{\theta_3}}K^3q^{-1}+K^3q^{-1}e^{-CR^2/q}+CK^6q^{1/2}R^{-3/2}\le C\exp(-K^{\theta_3/3}), \qquad (7.32)$$

for sufficiently large K, and this proves (7.24) with $\theta=\theta_3/3$.

We will now prove (7.25). We will consider only the set $\mathcal{Q}^c_{\sigma,Z}$ and only for $Z=0$. The modification needed for the general case is only notational. We start the proof by noting that

$$\begin{aligned}&\frac{1}{1+s'}\int_0^{s'}\mathrm{d}a\frac{1}{M'}\sum_{i=-M'}^{M'}\frac{1}{|x_i(\sigma-a)-x_{i+1}(\sigma-a)|^2} \qquad (7.33)\\ &\qquad\le C\frac{1}{1+s}\int_0^s\mathrm{d}a\frac{1}{M}\sum_{i=-M}^{M}\frac{1}{|x_i(\sigma-a)-x_{i+1}(\sigma-a)|^2}\end{aligned}$$

holds for any $s' \in [s/2, s]$ and $M' \in [M/2, M]$. Hence it is enough to estimate the probability

$$\mathbb{P}^\omega\Big\{\frac{1}{1+s}\int_0^s \mathrm{d}a \frac{1}{M}\sum_{i=-M}^{M}\frac{1}{|x_i(\sigma-a)-x_{i+1}(\sigma-a)|^2} \geq K^\rho\Big\} \tag{7.34}$$

for fixed dyadic points $(s, M) = \{(2^{-p}K^2, 2^{-q}K)\}$ in space-time for each integer $p, q \leq C\log K$. Since the cardinality of the set of these dyadic points is just $C(\log K)^2$, it suffices to estimate (7.34) only for a fixed s, M.

The proof is different for $\beta = 1$ and $\beta > 1$. In the latter case, from (6.63) we see that the random variable in (7.34) has expectation $CK^{C_3\xi}$. Thus the probability in (7.34) is bounded by $CK^{C_3\xi-\varrho}$, so (7.25) holds in this case with C_4 slightly larger than C_3+1 to accommodate the $C(\log K)^2$ factor.

The case $\beta = 1$ is more complicated since the random variable in (7.34) has a divergent expectation. We will use stochastic continuity. The proof below would work for any $\beta > 1$ as well, but for simplicity we set $\beta = 1$. For any fixed time s_k from the set (7.29), for any $i \in [\![-K-1, K]\!]$ and for any $\delta > 0$, we have

$$\mathbb{P}^\omega\Big\{|x_i(s_k) - x_{i+1}(s_k)| \leq 3\delta\Big\} \leq C(N\delta)^2$$

from the weak form of level repulsion (6.58). Thus

$$\begin{aligned}\mathbb{P}^\omega\Big\{|x_i(s_k)-x_{i+1}(s_k)| \leq 3\delta \ : \ \forall\, i \in [\![-K-1, K]\!],\ \forall k = 0, 1, \ldots K^2/q\Big\} \\ \leq \sum_i\sum_k \mathbb{P}^\omega\Big\{|x_i(s_k) - x_{i+1}(s_k)| \leq 3\delta\Big\} \leq CK^3q^{-1}(N\delta)^2 \leq N^5\delta^2.\end{aligned} \tag{7.35}$$

Set

$$\Omega := \Big\{\sup_{0\leq t\leq C_1K\log K, |i|\leq K} |x_i(t) - x_{i+1}(t)| \geq \delta\Big\}.$$

Combining (7.35) with the stochastic continuity (7.30) (with $R = \delta$), we get

$$\mathbb{P}^\omega(\Omega^c) \leq CN^5\delta^2 + K^3q^{-1}e^{-C\delta^2/q} + CK^6q^{1/2}\delta^{-3/2}.$$

Choosing $\delta := \exp(-K^\xi)$, $q := \delta^4$, we obtain

$$\mathbb{P}^\omega(\Omega^c) \leq \exp(-cK^\xi). \tag{7.36}$$

On the set Ω we can estimate (7.34) by Markov inequality. By the strong level repulsion (6.60), we have

$$\begin{aligned}&\mathbb{E}\frac{\mathbf{1}(\Omega)}{1+s}\int_0^s \mathrm{d}a\frac{1}{M}\sum_{i=-M}^{M}\frac{1}{|x_i(\sigma-a)-x_{i+1}(\sigma-a)|^2} \\ &\qquad \leq C\sup_{0\leq t\leq s}\mathbb{E}^\omega\frac{\mathbf{1}(\Omega)}{|x_i(t)-x_{i+1}(t)|^2} \leq K^{2C_3\xi}\int_\delta^\infty \frac{1}{\ell^2}\,\ell\mathrm{d}\ell \leq CK^{2C_3\xi}|\log\delta| \leq CK^{(2C_3+1)\xi}\end{aligned}$$

(notice that the condition $\delta \geq \exp(-K^\theta)$ set for (6.60) is satisfied for $\theta = \xi$). Therefore

$$(7.34) \leq \mathbb{P}^\omega(\Omega^c) + CK^{(2C_3+1)\xi-\rho} \leq CK^{(2C_3+1)\xi-\rho}.$$

Choosing C_4 slightly larger than $2C_3+1$ (to accommodate the $C(\log K)^2$ factor from the dyadic argument), we have proved (7.25). □

7.2.1 Restriction to the set $\mathcal{G}$

Now we show that the expectation (7.19) can be restricted to the good set $\mathcal{G}$ with a small error. With a slight abuse of notations we use $\mathcal{G}$ also to denote the characteristic function of the set $\mathcal{G}$. We just estimate the complement as

$$\int \big|O'(x_p - x_{p+1})\big|\mathbb{E}_{\mathbf{x}}\mathcal{G}^c[w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)]\omega(\mathrm{d}\mathbf{x})$$
$$\leq C\int \mathbb{E}_{\mathbf{x}}\mathcal{G}^c\big[|w_p(\sigma,\mathbf{x}(\cdot);\sigma)| + |w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)|\big]\omega(\mathrm{d}\mathbf{x}).$$

Since $\mathcal{A}\geq 0$ as a $\mathcal{K}\times\mathcal{K}$ matrix, the equation (7.17) is contraction in L^2. Clearly $\mathcal{A}$ is a contraction in L^1 as well, hence it is a contraction in any L^q, $1\leq q\leq 2$, by interpolation. By the Schwarz inequality and the L^q-contraction for some $1<q<2$, we have

$$\begin{aligned}\mathbb{E}^\omega\mathcal{G}^c|w_p(\sigma,\mathbf{x}(\cdot);\sigma)| &\leq \big[\mathbb{E}^\omega\mathcal{G}^c\big]^{q/(q-1)}\big[\mathbb{E}^\omega|w_p(\sigma,\mathbf{x}(\cdot);\sigma)|^q\big]^{1/q}\\ &\leq \big[\mathbb{P}^\omega\mathcal{G}^c\big]^{q/(q-1)}\Big[\mathbb{E}^\omega\sum_{i\in I}|w_i(0,\mathbf{x}(\cdot);\sigma)|^q\Big]^{1/q}\\ &\leq CK^{1+C_3\xi}e^{-cK^{\theta_4}}\leq e^{-cK^{\theta_4}}\end{aligned}$$

with some $\theta_4>0$. Here we used (7.24) for the first factor. For the second factor we recall $\mathbf{w}(0,\mathbf{x}(\cdot);\sigma)=\nabla h_0(\mathbf{x}(\sigma))$, which can be estimated by (6.63) using the invariance of the measure ω (recall the definition of h_0 from (6.52)).

Hence we have proved that

$$\int \big|O'(x_p - x_{p+1})\big|\mathbb{E}_{\mathbf{x}}\mathcal{G}^c[w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)]\omega(\mathrm{d}\mathbf{x})\leq Ce^{-cK^{\theta_4}}.$$

Therefore, under the conditions of Corollary 7.1, and using the notation $\mathbb{E}^\omega$ for the process, we have

$$|\langle h_0; O(x_p - x_{p+1})\rangle_\omega| \leq \int_0^{C_1K\log K}\mathrm{d}\sigma\mathbb{E}^\omega\mathcal{G}\big|w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)\big| + O(K^{-2}), \quad (7.37)$$

where $\mathbf{w}$ is the solution to (7.17) with initial condition $\mathbf{w}_0=\nabla h_0(\mathbf{x}(\sigma))$, assuming that the constant C_1 in the upper limit of the integration is large enough.

7.2.2 Restriction to the set $\widetilde{\mathcal{Q}}$ and the decay estimates

The complement of the set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ includes the "bad" paths for which the level repulsion estimate in an average sense does not hold. However, the probability of $\widetilde{\mathcal{Q}}^c_{\sigma,Z}$ is not very small, it is only a small negative power of K, see (7.25). This estimate would not be sufficient against the time integration of order $C_1K\log K$ in (7.37); we will have to use an L^1-L^∞ decay property of (7.17) which we now derive.

Denote the L^p-norm of a vector $\mathbf{u}=\{u_j\ :\ j\in I\}$ by

$$\|\mathbf{u}\|_p=\Big(\sum_{j\in I}|u_j|^p\Big)^{1/p}. \quad (7.38)$$

We have the following decay estimate for the equation (7.17):

Proposition 7.4 *Consider the evolution equation*

$$\partial_s \mathbf{u}(s) = -\mathcal{A}(s)\mathbf{u}(s), \qquad \mathbf{u}(s) \in \mathbb{R}^I = \mathbb{R}^{\mathcal{K}} \tag{7.39}$$

and fix $\sigma > 0$. Suppose that for some constant b we have

$$B_{jk}(s) \geq \frac{b}{(j-k)^2}, \quad 0 \leq s \leq \sigma, \quad j \neq k, \tag{7.40}$$

and

$$W_j(s) \geq \frac{b}{d_j}, \qquad d_j := \big||j| - K\big| + 1, \quad 0 \leq s \leq \sigma. \tag{7.41}$$

Then for any $1 \leq p \leq q \leq \infty$ we have the decay estimate

$$\|\mathbf{u}(s)\|_q \leq (sb)^{-(\frac{1}{p}-\frac{1}{q})}\|\mathbf{u}(0)\|_p, \qquad 0 < s \leq \sigma. \tag{7.42}$$

Proof. We consider only the case $b = 1$, the general case follows from scaling. We follow the idea of Nash and start from the L^2-identity

$$\partial_s \|\mathbf{u}(s)\|_2^2 = -2\mathfrak{a}(s)[\mathbf{u}(s), \mathbf{u}(s)]. \tag{7.43}$$

For each s we can extend $\mathbf{u}(s) : I \to \mathbb{R}^{\mathcal{K}}$ to a function $\widetilde{\mathbf{u}}(s)$: on $\mathbb{Z}$ by defining $\widetilde{u}_j(s) = u_j(s)$ for $|j| \leq K$ and $\widetilde{u}_j(s) = 0$ for $j > |K|$. Dropping the time argument, we have, by the estimates (7.40) and (7.41) with $b = 1$,

$$2\mathfrak{a}(\mathbf{u}, \mathbf{u}) \geq \sum_{i,j\in\mathbb{Z}} \frac{(\widetilde{u}_i - \widetilde{u}_j)^2}{(i-j)^2} \geq c\|\widetilde{\mathbf{u}}\|_4^4\, \|\widetilde{\mathbf{u}}\|_2^{-2}, \tag{7.44}$$

with some positive constant, where, in the second step, we used the Gagliardo-Nirenberg inequality for the discrete operator $|p|$, see (B.1) in the Appendix. Thus we have

$$\mathfrak{a}[\mathbf{u}, \mathbf{u}] \geq c\|\mathbf{u}\|_4^4\, \|\mathbf{u}\|_2^{-2}, \tag{7.45}$$

and the energy inequality

$$\partial_s\|\mathbf{u}\|_2^2 \leq -c\|\mathbf{u}\|_4^4\, \|\mathbf{u}\|_2^{-2} \leq -c\|\mathbf{u}\|_2^4\, \|\mathbf{u}\|_1^{-2}, \tag{7.46}$$

using the Hölder estimate $\|\mathbf{u}\|_2 \leq \|\mathbf{u}\|_1^{1/3}\|\mathbf{u}\|_4^{2/3}$. Integrating this inequality from 0 to s we get

$$\|\mathbf{u}(s)\|_2 \leq Cs^{-1/2}\|\mathbf{u}(0)\|_1, \tag{7.47}$$

and similarly we also have $\|\mathbf{u}(2s)\|_2 \leq Cs^{-1/2}\|\mathbf{u}(s)\|_1$. Since the previous proof uses only the time independent lower bounds (7.40), (7.41), we can use duality in the time interval $[s, 2s]$ to have

$$\|\mathbf{u}(2s)\|_\infty \leq Cs^{-1/2}\|\mathbf{u}(s)\|_2.$$

Together with (7.47) we have

$$\|\mathbf{u}(2s)\|_\infty \leq Cs^{-1}\|\mathbf{u}(0)\|_1.$$

By interpolation, we have thus proved the Lemma. $\square$

In the good set $\mathcal{G}$ (see (7.21)), the bounds (7.40) and (7.41) hold with $b = cK^{-\xi'}$. Hence from the decay estimate (7.42), for any fixed σ, Z, we can insert the other good set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ into the expectation in (7.37). This is obvious since the contribution of its complement is bounded by

$$
\begin{aligned}
\int_0^{C_1 K\log K} & \mathrm{d}\sigma\, \mathbb{E}^\omega\, \widetilde{\mathcal{Q}}^c_{\sigma,Z}\mathcal{G}\big|w_p(\sigma,\mathbf{x}(\cdot);\sigma) - w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)\big| \\
& \le CK^{\xi'}\int_0^{C_1 K\log K}\mathrm{d}\sigma\ \sigma^{-\frac{1}{1+\xi}}\ \mathbb{E}^\omega\Big[\mathcal{G}\big(\sum_{i\in I}|\nabla_i h_0(\mathbf{x}(\sigma))|^{1+\xi}\big)^{\frac{1}{1+\xi}}\,\widetilde{\mathcal{Q}}^c_{\sigma,Z}\Big] \\
& \le CK^{2\xi'}\int_0^{C_1 K\log K}\mathrm{d}\sigma\ \sigma^{-\frac{1}{1+\xi}}\ \mathbb{E}^\omega\Big[\big[1+d(x_K(\sigma))^{-1}+d(x_{-K}(\sigma))^{-1}\big]\widetilde{\mathcal{Q}}^c_{\sigma,Z}\Big] \\
& \le CK^{2\xi'}\int_0^{C_1 K\log K}\mathrm{d}\sigma\ \sigma^{-\frac{1}{1+\xi}}\ \Big[\mathbb{E}^\omega\big[1+d(x_K(\sigma))^{-1}+d(x_{-K}(\sigma))^{-1}\big]^{3/2}\Big]^{2/3}\Big[\mathbb{P}^\omega\big(\widetilde{\mathcal{Q}}^c_{\sigma,Z}\big)\Big]^{\frac{1}{3}} \\
& \le CK^{2\xi'}(C_1K\log K)^{\xi}K^{C_3\xi}K^{(C_4\xi-\rho)/3}.
\end{aligned}
\tag{7.48}
$$

where we used the decay estimate (7.42) with $q=\infty$, $p=1+\xi$ in the second line. In the third line we split the sum into two parts and used the bound

$$
|\partial_j h_0(\mathbf{x})| \le |(V_{\mathbf{y}})'(x_j) - (\widetilde{V}_{\widetilde{\mathbf{y}}})'(x_j)| \le \frac{K^{\xi'}}{d(x_j)},
\tag{7.49}
$$

that follows from (6.42) (with ξ replaced by ξ^2 since $\mathbf{y},\widetilde{\mathbf{y}}\in\mathcal{R}_{L,K}(\xi^2\delta/2,\alpha/2)$). Recall that $d(x)$ is the distance to the boundary, see (6.44). For indices away from the boundary, $|i|\le K-CK^{\xi'}$, we have $|d(x_i)|\ge K^{-\xi'}\min\{|i-K|,|i+K|\}$ on the set $\mathcal{G}$ that guarantees the finiteness of the sum. For indices near the boundary we just used the worst term with $i=\pm K$. We used a Hölder inequality in the fourth line of (7.48) and computed the expectation by using stationarity and (6.63) in the last line. Hence we have proved the following proposition:

Proposition 7.5 *For any $\sigma>0$ and for any path $\{\mathbf{x}(\cdot):[0,\sigma]\to\mathbb{R}^{\mathcal{K}}\}$ let $\mathbf{w}$ be the solution of* (7.17) *with h_0 given by* (6.52). *Suppose that*

$$
\rho \ge 12\xi' + 6(C_4+C_3+1)\xi
\tag{7.50}
$$

holds with C_4 defined in (7.25). *Then for any fixed $Z, p\in I$ with $p\neq K$, we have*

$$
|\langle h_0; O(x_p-x_{p+1})\rangle_\omega| \le \int_0^{C_1K\log K}\mathrm{d}\sigma\mathbb{E}^\omega\,\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\big|w_p(\sigma,\mathbf{x}(\cdot);\sigma)-w_{p+1}(\sigma,\mathbf{x}(\cdot);\sigma)\big| + O(K^{-\rho/6}).
\tag{7.51}
$$

7.3 Short time cutoff and finite speed of propagation

The Hölder continuity of the parabolic equation (7.17) emerges only after a certain time, thus for the small σ regime in the integral (7.51) we need a different argument. Since we are interested in the Hölder continuity around the middle of the interval I (note that $|p|\le K^{1-\xi^*}$ in Theorem 6.5), and the initial condition ∇h_0 is small in this region, a finite speed of propagation estimate will guarantee that $w_p(\sigma;\mathbf{x}(0),\sigma)$ is small if σ is not too large.

From now on, we fix $\sigma\le C_1K\log K$, $|Z|\le K/2$ and a path $\mathbf{x}(\cdot)$, and assume that $\mathbf{x}(\cdot)\in\mathcal{G}\cap\widetilde{\mathcal{Q}}_{\sigma,Z}$. In particular, thanks to the definition of $\mathcal{G}$ and the regularity of the locations α_j, the

time dependent coefficients $B_{ij}(s)$ and $W_i(s)$ of the equation (7.17) satisfy (7.40) and (7.41) with $b = K^{-\xi'}$.

Let $\mathbf{q}(s) := \mathbf{w}(s; \mathbf{x}(\cdot), \sigma)$ which satisfies the equation (7.17) with initial data $\mathbf{q}(0) = \nabla h_0(\mathbf{x}(\sigma))$, where h_0 is defined in (6.52). We also recall the bound (7.49). We split the solution $\mathbf{q} = \mathbf{q}^{in} + \mathbf{q}^{out}$ according to the support of the initial condition, i.e. $\mathbf{q}^{in}$ and $\mathbf{q}^{out}$ are solutions to (7.17) with initial data

$$q_j^{in}(0) := \partial_j h_0(\mathbf{x}(\sigma))\mathbf{1}(|j| \le K^{1-\theta_5}), \quad q_j^{out}(0) := \partial_j h_0(\mathbf{x}(\sigma))\mathbf{1}(|j| > K^{1-\theta_5}) \tag{7.52}$$

with some small parameter $\theta_5 > 0$. From (7.49) and the rigidity bound given by $\mathcal{G}$, we have, on the set $\mathcal{G}$, that

$$\left\|\mathbf{q}^{in}(0)\right\|_1 \le K^{\xi' - \theta_5}, \tag{7.53}$$

From the $L^1 - L^\infty$ decay estimate (7.42), we have

$$\int_0^{C_1 K \log K} \mathrm{d}\sigma \; \mathbb{E}^\omega \mathcal{G}\left|q_p^{in}(\sigma) - q_{p+1}^{in}(\sigma)\right| \le K^{\xi' - \theta_5} \int_0^{C_1 K \log K} \sigma^{-1} \, \mathrm{d}\sigma \le K^{2\xi' - \theta_5}. \tag{7.54}$$

(Recall that the singularity $\sigma \sim 0$ can be cutoff as in (7.48)). Together with (7.51) and with the choice

$$\theta_5 > \rho \tag{7.55}$$

and recalling $\rho \ge 4\xi'$ from (7.50), we have

$$|\langle h_0; O(x_p - x_{p+1})\rangle_\omega| \le \int_0^{C_1 K \log K} \mathrm{d}\sigma \; \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma, Z} \mathcal{G}\left|q_p^{out}(\sigma) - q_{p+1}^{out}(\sigma)\right| + O(K^{-\rho/6}). \tag{7.56}$$

It remains to control q^{out}. In order to study this question, we consider the fundamental solution of the equation (7.17), i.e., for a fixed, the solution

$$\partial_s \mathbf{u}^a(s) = -\mathcal{A}(s)\mathbf{u}^a(s), \quad u_j^a(0) = \delta_{aj}. \tag{7.57}$$

Note that the equation is the same as the one studied in (7.17), but the initial condition is different; in particular previously we studied this equation with a random initial data. By linear superposition we then have

$$q_p^{out}(t) = \sum_{|a| > K^{1-\theta_5}} \partial_a h_0(\mathbf{x}(\sigma)) u_p^a(t) \tag{7.58}$$

for any $t > 0$.

We will assume that the coefficients of $\mathcal{A}$ satisfy for some fixed $|Z| \le K/2$ and $\rho_1 > 0$ the bound

$$\sup_{0 \le s \le \sigma} \sup_{0 \le M \le K} \frac{1}{1+s} \int_0^s \frac{1}{M} \sum_{i \in I \,:\, |i-Z| \le M} \; \sum_{j \in I \,:\, |j-Z| \le M} B_{ij}(\sigma - s) \mathrm{d}s \le CK^{\rho_1}, \tag{7.59}$$

and the similar bound if Z is replaced with K and $-K$. Notice that on the set $\mathcal{G} \cap \mathcal{Q}_{\sigma,Z}$ and with the choice

$$\rho_1 := \rho + \xi',$$

(7.59) follows from (7.22), since for the summands with $|i - j| \ge K^{\xi'}$ in (7.59) we can use $B_{ij} \le C|\alpha_i - \alpha_j|^{-2} \le C|i-j|^{-2}$.

The following lemma provides a finite speed of propagation estimate for the equation (7.57). This estimate is not optimal, but it is sufficient for our purpose. The proof will be given in the next section.

Lemma 7.6 *[Finite Speed of Propagation Estimate] Fix $a \in I$ and $\sigma \le C_1 K \log K$. Denote by $\mathbf{u}(s) = \mathbf{u}^a(s)$ the solution to (7.57) with initial condition $u_j^a(0) = \delta_{aj}$. We assume that the coefficients of $\mathcal{A}$ satisfy*

$$W_i(s) \ge \frac{K^{-\xi'}}{d_i}, \qquad B_{ij}(s) \ge \frac{K^{-\xi'}}{|i-j|^2}, \qquad 0 \le s \le \sigma, \tag{7.60}$$

where $d_i := \min\{|i+K|, |i-K|\}$. Assume that (7.59) is satisfied for some fixed Z, $|Z| \le K/2$. Then we have the estimate for any $s \le \sigma$

$$|u_p^a(s)| \le \frac{CK^{\rho_1 + 2\xi' + 1/2}\sqrt{s+1}}{|p-a|}. \tag{7.61}$$

7.4 Proof of the Finite Speed of Propagation Estimate, Lemma 7.6

Let $1 \ll \ell \ll K$ be a parameter to be specified later. Split the time dependent operator $\mathcal{A} = \mathcal{A}(s)$ defined in (7.15) into a short range and a long range part, $\mathcal{A} = \mathcal{S} + \mathcal{R}$, with

$$(\mathcal{S}\mathbf{v})_j := - \sum_{k\,:\,|j-k|\le \ell} B_{jk}(v_k - v_j) + W_j v_j \tag{7.62}$$

and

$$(\mathcal{R}\mathbf{v})_j := - \sum_{k\,:\,|j-k|>\ell} B_{jk}(v_k - v_j). \tag{7.63}$$

Note that $\mathcal{S}$ and $\mathcal{R}$ are time dependent. Denote by $U_{\mathcal{S}}(s_1, s_2)$ the semigroup associated with $\mathcal{S}$ from time s_1 to time s_2, i.e.

$$\partial_{s_2} U_{\mathcal{S}}(s_1, s_2) = -\mathcal{S}(s_2) U_{\mathcal{S}}(s_1, s_2)$$

for any $s_1 \le s_2$, and $U_{\mathcal{S}}(s_1, s_1) = I$; the notation $U_{\mathcal{A}}(s_1, s_2)$ is analogous. Then by the Duhamel formula

$$\mathbf{u}(s) = U_{\mathcal{S}}(0, s)\mathbf{u}_0 + \int_0^s U_{\mathcal{A}}(s', s)\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{u}_0 \mathrm{d}s'. \tag{7.64}$$

Notice that for $\ell \gg K^{\xi'}$ and for $\mathbf{x}(\cdot)$ in the good set $\mathcal{G}$ (see (7.21)), we have

$$\|\mathcal{R}\mathbf{v}\|_1 = \sum_{|j|\le K} \Big| \sum_{k:|j-k|\ge \ell} \frac{1}{(x_j - x_k)^2} v_k \Big| \le C\ell^{-1}\|\mathbf{v}\|_1, \tag{7.65}$$

or more generally,

$$\|\mathcal{R}\mathbf{v}\|_p \le C\ell^{-1}\|\mathbf{v}\|_p, \quad 1 \le p \le \infty, \qquad \|\mathcal{R}\mathbf{v}\|_\infty \le C\ell^{-2}\|\mathbf{v}\|_1. \tag{7.66}$$

Recall the decay estimate (7.42) for the semigroup $U_{\mathcal{A}}$ that is applicable by (7.60). Hence we have, for $s \ge 2$,

$$\int_0^s \|U_{\mathcal{A}}(s', s)\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{u}_0\|_\infty \,\mathrm{d}s'$$
$$\le K^{\xi'} \int_0^s (s-s')^{-1} \|\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{u}_0\|_1 \,\mathrm{d}s' \le K^{\xi'}\ell^{-1}(\log s)\|\mathbf{u}_0\|_1,$$

where we used that $U_{\mathcal{S}}$ is a contraction on L^1. The non-integrable short time singularity for s' very close to s, $|s-s'|\le K^{-C}$, can be removed by using the $L^p\to L^\infty$ bound (7.42) with some $p>1$, invoking a similar argument in (7.48). In this short time cutoff argument we used that $U_{\mathcal{S}}$ is an L^p contraction for any $1\le p\le 2$ by interpolation, and that the rate of the $L^p\to L^\infty$ decay of $U_{\mathcal{A}}$ are given in (7.42).

$$\|\mathbf{u}(s)-U_{\mathcal{S}}(0,s)\mathbf{u}_0\|_\infty \le \ell^{-1}(\log s)K^{\xi'} \le C\ell^{-1}(\log K)K^{\xi'}, \tag{7.67}$$

where we have used that $\mathbf{x}(\cdot)$ is in the good set $\mathcal{G}$ and that $s\le C_1K\log K$.

We now prove a cutoff estimate for the short range dynamics. Let $\mathbf{r}(s):=U_{\mathcal{S}}(0,s)\mathbf{u}_0$ and define

$$f(s)=\sum_j \phi_j r_j^2(s), \quad \phi_j=e^{|j-a|/\theta} \tag{7.68}$$

with some parameter $\theta\ge\ell$ to be specified later. Recall that a is the location of the initial condition, $\mathbf{u}_0=\delta_a$. In particular, $f(0)=1$.

Differentiating f and using $W_j\ge 0$, we have

$$\begin{aligned} f'(s)=\partial_s\sum_j\phi_j r_j^2(s) &\le 2\sum_j\phi_j\sum_{k:|j-k|\le\ell} r_j(s)B_{kj}(s)(r_k-r_j)(s)\\ &=\sum_{|j-k|\le\ell} B_{kj}(s)(r_k-r_j)(s)\left[r_j(s)\phi_j-r_k(s)\phi_k\right]\\ &=\sum_{|j-k|\le\ell} B_{kj}(s)(r_k-r_j)(s)\phi_j\left[r_j-r_k\right](s)\\ &\quad+\sum_{|j-k|\le\ell} B_{kj}(s)(r_k-r_j)(s)\left[\phi_j-\phi_k\right]r_k(s). \end{aligned}$$

In the second term we use Schwarz inequality and absorb the quadratic term in r_k-r_j into the first term that is negative. Assuming $\ell\le\theta$, we have $\phi_k^{-2}\left[\phi_j-\phi_k\right]^2\le C\ell^2/\theta^2$ for $|j-k|\le\ell$. Thus

$$f'(s)\le C\sum_{|j-k|\le\ell}B_{kj}(s)\phi_k^{-1}\left[\phi_j-\phi_k\right]^2 r_k^2(s)\le C\theta^{-2}\ell^2\left(\sum_{k',j:|j-k'|\le\ell}B_{k'j}(s)\right)\sum_k\phi_k r_k^2(s).$$

From a Gromwall argument we have

$$f(s)\le\exp\left[C\theta^{-2}\ell^2\int_0^s\sum_{k,j:|j-k|\le\ell}B_{kj}(s')\mathrm{d}s'\right]f(0). \tag{7.69}$$

From the assumption (7.59) with $M=K$ and any arbitrary Z, we can bound the integration in the exponent by

$$\int_0^s\sum_{k,j:|j-k|\le K}B_{kj}(s')\mathrm{d}s'\le K^{1+\rho_1}(s+1). \tag{7.70}$$

Thus we have

$$\sum_j e^{|j-a|/\theta}r_j^2(s)=f(s)\le\exp\left[\theta^{-2}\ell^2K^{1+\rho_1}(s+1)\right]f(0)\le C, \tag{7.71}$$

provided that we choose

$$\theta = \ell K^{(\rho_1+1)/2}\sqrt{s+1}. \tag{7.72}$$

Now we choose

$$\ell = |p-a|K^{-\xi'-(\rho_1+1)/2}(s+1)^{-1/2}$$

so that $e^{|p-a|/\theta} \geq \exp\left(K^{\xi'}\right)$. Using this choice in (7.71) and (7.67) to estimate $u_p(s)$, we have thus proved that

$$|u_p(s)| \leq \ell^{-1}(\log K)K^{\xi'} + Ce^{-K^{-\xi'}} \leq \frac{K^{2\xi'+(\rho_1+1)/2}\sqrt{s+1}}{|p-a|}. \tag{7.73}$$

This concludes the proof of Lemma 7.6.

7.5 Completing the proof of Theorem 6.5

In this section we complete the proof of Theorem 6.5 assuming a discrete version of the De Giorgi-Nash-Moser Hölder regularity estimate for parabolic equations with singular coefficients (Theorem 7.7 below).

Notice that on the set $\mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,Z}$ the conditions of Lemma 7.6 are satisfied, especially (7.59) follows from the definition (7.22) as we already remarked. Thus we can use (7.61) and (7.58) to estimate the short time integration regime in (7.56). Setting

$$\theta_5 := \min\left\{\frac{\xi^*}{2}, \frac{1}{100}\right\}, \tag{7.74}$$

we obtain, for any $|Z| \leq 2K^{1-\xi^*}$ and $|p| \leq K^{1-\xi^*}$,

$$\begin{aligned}
&\int_0^{K^{1/4}} \mathrm{d}\sigma\, \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\left|q_p^{out}(\sigma) - q_{p+1}^{out}(\sigma)\right| \\
&\quad \leq C\int_0^{K^{1/4}} \mathrm{d}\sigma\, \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G} \sum_{|a|>K^{1-\theta_5}} |\partial_a h_0(\mathbf{x}(\sigma))| u_p^a(\sigma) \\
&\quad \leq CK^{2\xi'+\rho_1+1/2+\frac{1}{4}+\frac{1}{8}-(1-\theta_5)} \sup_{1\leq\sigma\leq K^{1/4}} \mathbb{E}^\omega\, \widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\int_0^1 \mathrm{d}s \sum_{|a|>K^{1-\theta_5}} |\partial_a h_0(\mathbf{x}(\sigma-s))| \\
&\quad \leq CK^{4\xi'+\rho_1+\theta_5-\frac{1}{8}} \sup_{1\leq\sigma\leq K^{1/4}} \mathbb{E}^\omega\, \widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\int_0^1 \mathrm{d}s\Big[\frac{1}{d(x_K(\sigma-s))} + \frac{1}{d(x_{-K}(\sigma-s))}\Big] \\
&\quad \leq CK^{4\xi'+\rho_1+C_3\xi+\theta_5-\frac{1}{8}} \leq K^{-\frac{1}{10}}
\end{aligned} \tag{7.75}$$

provided that

$$4\xi' + \rho_1 + C_3\xi \leq \frac{1}{100}. \tag{7.76}$$

In the third line above we used (7.61) together with $|p-a| \geq \frac{1}{2}K^{1-\theta_5}$. This latter bound follows from $|a| > K^{1-\theta_5}$ and $|p| \leq K^{1-\xi^*}$ and from the choice $\theta_5 < \xi^*$. In the fourth line we used (7.49) and that on the set $\mathcal{G}$ we have

$$\sum_j \frac{1}{d(x_j)} \leq (\log K)K^{\xi'}\Big[\frac{1}{d(x_K)} + \frac{1}{d(x_{-K})}\Big].$$

Moreover, in the last step we used (6.63).

We now treat the integration regime $\sigma \in [K^{1/4}, C_1 K \log K]$ in (7.56) which requires a partial regularity result which we now formulate. We define the regular kernel by

$$\widetilde{B}_{ij} = \frac{\varrho(0)^2}{(i-j)^2}, \tag{7.77}$$

where we recall that $\varrho(0)$ is the equilibrium density at the center. Since $\mathbf{x}(\cdot) \in \mathcal{G}$, we have

$$\sup_{s \le C_1 K \log K} \sup_{|i-j| \ge \widehat{C} K^{\xi'}} \left| B_{ij}(s) - \widetilde{B}_{ij} \right| \le K^{-\xi'} \widetilde{B}_{ij} \tag{7.78}$$

for some sufficiently large $\widehat{C}$ depending on $\varrho(0)$. In particular, there is a constant C such that

$$\frac{1}{C(i-j)^2} \le B_{ij}(s) \le \frac{C}{(i-j)^2} \tag{7.79}$$

for any $|i-j| \ge \widehat{C} K^{\xi'}$ and $0 \le s \le C_1 K \log K$.

Theorem 7.7 (Parabolic partial regularity with singular coefficients) *Let $\mathbf{u}$ be a solution to (7.57), where $\mathbf{u} = \mathbf{u}^a$ for any choice of a. Suppose that there exist ξ_0, ρ_0 small enough such that for some $\xi' \le \xi_0$ and $\rho_1 \le \rho_0$ the coefficients of $\mathcal{A}$ satisfy (7.60), (7.59), (7.79) and*

$$W_i \le \frac{K^{\xi'}}{d_i}, \quad \textit{if} \quad d_i \ge K^{C\xi'}. \tag{7.80}$$

Let $\sigma \in [K^{c_3}, C_1 K \log K]$ be fixed, where $c_3 > 0$ is an arbitrary positive constant. Then for any $0 < q' < 1$ there exists $q > 0$ depending only on ξ_0, ρ_0 and q' so that for any $|Z| \le K/2$

$$\sup_{\max(|j-Z|, |j'-Z|) \le \sigma^{1-q'}} |u_j(\sigma) - u_{j'}(\sigma)| \le C\sigma^{-1-q}, \tag{7.81}$$

where $\mathbf{u} = \mathbf{u}^a$ for any choice of a. The constant C in (7.81) depends only on the control parameters, c_3, ρ_0, ξ_0, but it is uniform in σ.

Notice that this result is deterministic. We also remark that if we define the rescaled function $v(j/K, t) := t u_j(t)$, then (7.81) can be interpreted as a type of Hölder regularity of v on scale $\sigma^{1-q'} K^{-1} \ll 1$ at the point Z/K:

$$|v(x,\sigma) - v(y,\sigma)| \le \sigma^{-q} \le |x-y|^{c_3 q} \tag{7.82}$$

for $1/K \le |x-y| \le \sigma^{1-q'}/K$ and x, y near Z/K. The Hölder exponent is thus at least $c_3 q$.

We have stated the regularity for $|Z| \le K/2$ to avoid dealing with the boundary regularity. For the application in the proof of Theorem 6.5, we need the regularity only for $Z = p$ (or any Z such that $|Z - p|$ is bounded). We will prove Theorem 7.7 in Sections 8.

Assuming Theorem 7.7, we now complete the proof of Theorem 6.5. As we already remarked, the conditions of Theorem 7.7 are satisfied on the set $\widetilde{\mathcal{Q}}_{\sigma,Z} \cap \mathcal{G}$ with some ρ_0, ξ_0 small universal

constants. For any $|p| \le K^{1-\xi^*}$ fixed, we choose $Z = p$ (in fact, we could choose any Z with $|Z-p| \le C$). Using (7.58) and (7.49), we have, for the large time integration regime in (7.56),

$$\begin{aligned}
&\int_{K^{1/4}}^{C_1 K \log K} \mathrm{d}\sigma \, \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z} \mathcal{G} \big| q_p^{out}(\sigma) - q_{p+1}^{out}(\sigma) \big| \qquad (7.83) \\
&\qquad \le CK^{\xi'} \int_{K^{1/4}}^{C_1 K \log K} \mathrm{d}\sigma \, \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z} \mathcal{G} \sum_{|a|>K^{1-\theta_5}} \frac{1}{d(x_a(\sigma))} \big| u_p^a(\sigma) - u_{p+1}^a(\sigma) \big| \\
&\qquad \le CK^{2\xi'} \int_{K^{1/4}}^{C_1 K \log K} \mathrm{d}\sigma \, \sigma^{-1-q} \, \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z} \mathcal{G} \int_0^1 \mathrm{d}s \sum_{|a|>K^{1-\theta_5}} \frac{1}{d(x_a(\sigma - s))} \\
&\qquad \le CK^{3\xi' + \rho_1 + C_3 \xi - \frac{1}{4} q}.
\end{aligned}$$

In the third line we used Theorem 7.7 with $c_3 = 1/4$ and $q' = 1/2$, which guarantees that $|Z - p| \le C \le \sigma^{1-q'}$ for any $\sigma \ge K^{1/4}$. The exponent q, corresponding to the choice $q' = 1/2$, is a small universal constant. In the last line of (7.83) we used a similar argument as in the last step of (7.75).

Finally, from (7.56), (7.75) and (7.83) and $\rho_1 = \rho + \xi'$ we have

$$|\langle h_0; O(x_p - x_{p+1}) \rangle_\omega| \le CK^{4\xi' + \rho + C_3 \xi - \frac{1}{4} q} + O(K^{-\frac{1}{10}}) + O(K^{-\rho/6}). \qquad (7.84)$$

For a given $\xi^* > 0$, recall that we defined $\theta_5 := \min\{\xi^*/2, \frac{1}{100}\}$ and we now choose

$$\rho := \min\Big\{ \frac{q}{10}, \frac{\theta_5}{2} \Big\} = \min\Big\{ \frac{q}{10}, \frac{\xi^*}{4}, \frac{1}{200} \Big\}, \qquad (7.85)$$

in particular (7.55) is satisfied. It is then clear that for any sufficiently small ξ all conditions in (7.76) and (7.50) on the exponents ξ, $\xi' = (C_2 + 1)\xi^2$ and $\rho_1 = \rho + \xi'$ can be simultaneously satisfied. Thus we can make the error term in (7.84) smaller than $K^{C\xi} K^{-\rho/6}$. With the choice of $\varepsilon = \rho/6$, where ρ is defined in (7.85), we proved Theorem 6.5. □

Although the choices of parameters seem to be complicated, the underlying mechanism is that there is a positive exponent q in (7.81), which to a great degree is an universal constant. This exponent provides an extra smallness factor in addition to the natural size of $u_j(\sigma)$, which is σ^{-1} from the $L^1 \to L^\infty$ decay. As (7.82) indicates, this gain comes from a Hölder regularity on the relevant scale. The parameters ξ, ξ' and ξ^* can be chosen arbitrarily small (without affecting the value of q). These parameters govern the cutoff levels in the regularization of the coefficients of $\mathcal{A}$. There are other minor considerations due to an additional cutoff for small time where we have to use a finite speed estimate. But the arguments for this part are of technical nature and most estimates are not optimized. We just worked out estimates sufficient for the purpose of proving Theorem 6.5. The choices of exponents related to the various cutoffs do not have intrinsic meanings.

As a guide to the reader, our choice of parameters, roughly speaking, are given by the following rule: We first fix a small parameter ξ^*. Then we choose the cutoff parameter θ_5 to be slightly smaller than ξ^*, (7.74). The exponent ρ in (7.22) has a lower bound by ξ and ξ' in (7.50). On the other hand, ρ will affect the cutoff bound and so we have the condition $\rho < \theta_5$ (i.e., (7.55)). So we choose $\rho \lesssim \xi^*$ and make ξ, ξ' very small so that the lower bound requirement on ρ is satisfied. Finally, if the parameter $\xi^* \le q/10$, we can use the gain from the Hölder continuity to compensate all the errors which depend only on ξ, ξ', ξ^*.

8 A discrete De Georgi-Nash-Moser estimate

In this section we prove Theorem 7.7, which is a Hölder continuity estimate for the parabolic evolution equation (7.57). Our equation is of the type considered in [10], but it is discrete and in a finite interval. The key difference, however, is that the coefficient $B_{ij} = (x_i - x_j)^{-2}$ in the elliptic part of (7.57) can be singular, while [10] assumed boundedness. In fact, by extending the reasoning of Ben Arous and Bourgade [3], the minimal gap $\min_i(x_{i+1} - x_i)$ for GOE is typically of order $N^{-1/2}$ in the microscopic coordinates we are using now. The only control we have for the singular behavior of B_{ij} is the estimate (7.59). This estimate essentially says that the space-time maximum function of $B_{i,i+1}(t)$ at a fixed space-time point $(Z, 0)$ is bounded by K^ϱ. Our main task is to show that this condition is sufficient for proving Hölder continuity at the same point. Our strategy follows closely the approach of Caffarelli-Chan-Vasseur [10] and the key cutoff functions (8.1, 8.30) are also the same as in [10]. The main new feature of our argument is the derivation of the local energy estimate, Theorem 8.1, for parabolic equation with singular coefficients satisfying (7.59) and (7.60). The proof of Theorem 8.1 will proceed in two steps. We first use the condition (7.59) and the argument of the energy estimate in [10] to provide a bound in $L^\infty_t(L^2(\mathbb{Z}))$ on the solution to (7.57) (Lemma 8.2). Using this estimate we run the argument again to improve it to an L^∞ estimate in space to obtain Theorem 8.1. Besides this proof, the derivation of the second De Giorgi estimate (Lemma 8.5) is also adjusted to the weaker condition (7.59). After that, starting from Section 8.2.2, the argument of [10] takes over and the proof presented here is the same as in [10]. Some more explanation will be given along the proofs.

We warn the reader that the notations of various constants in this section will follow [10] as much as possible for the sake of easy comparison with the paper [10]. The conventions of these constants will differ from the ones in the previous sections, in particular, we will restate all conditions.

8.1 Local dissipation estimate (first De Giorgi lemma)

Fix a large integer number M and a center $Z \in I$ with $M \le Z \le K - M$. Let $B = B_{M,Z} = \{i \in I \; |i - Z| \le M\}$ be the interval of length $2M$ about Z, and let it be equipped with the counting measure.

For any $\ell > 0$ define the function

$$\psi_i = \psi_i^{(M,Z,\ell)} := \ell\Big(\Big|\frac{i - Z}{M}\Big|^{1/2} - 1\Big)_+ \tag{8.1}$$

(this corresponds to the case $s = 1$, $N = 1$ with the notation of [10]). Note that this also has size of order ℓ and localizes on a scale M. Let

$$\psi_i^\ell = \ell + \psi_i.$$

Here ℓ will thus play the role of the typical size of $u - \psi$. One could scale out ℓ completely, but we keep it in. For any real number a, we will use the notation $a_+ = \max(a, 0) \ge 0$ and $a_- = \min(a, 0) \le 0$, in particular $a = a_+ + a_-$. In this section, we will use the convention that the time variable is in some interval $[-C_K, 0]$ where C_K is some constant depending on K. This convention is widely used for parabolic equations and in particular in [10]. Later on in our application, we will need to make an obvious shift in time.

Theorem 8.1 (Parabolic energy estimate with singular coefficients) *There exists a small positive constant ε_0 with the following properties. Let $\mathbf{u}$ be the solution to* (7.57) *and let M be*

defined by $M^\vartheta = K$ with some $\vartheta \geq 1$. Set $\sigma \leq C_1 K \log K$. We assume the following conditions on the matrix elements of $\mathcal{A} = \mathcal{B} + \mathcal{W}$:

i) For a fixed $|Z| \leq K/2$ and $\rho > 0$ assume that

$$\sup_{0\leq s\leq\sigma} \sup_{0\leq M\leq K} \frac{1}{1+s}\int_0^s \frac{1}{M} \sum_{i\in I\,:\,|i-Z|\leq M} \sum_{j\in I\,:\,|j-Z|\leq M} B_{ij}(\sigma - s)\mathrm{d}s \leq CK^\rho. \tag{8.2}$$

ii) For some $\xi > 0$ we have

$$W_i(s) \geq \frac{K^{-\xi}}{d_i}, \qquad B_{ij}(s) \geq \frac{K^{-\xi}}{|i-j|^2}, \qquad 0 \leq s \leq \sigma, \tag{8.3}$$

where $d_i := \min\{|i+K|, |i-K|\}$.

iii)

$$\frac{1}{C(i-j)^2} \leq B_{ij}(s) \leq \frac{C}{(i-j)^2} \tag{8.4}$$

for any $|i-j| \geq \widehat{C}K^\xi$ and $0 \leq s \leq \sigma$.

Suppose

$$\sup_{t\in[-2M,0]} \|u(t)\|_\infty \leq \ell M^{c_1} \tag{8.5}$$

for some c_1 sufficiently small such that

$$\Phi := 3c_1 + (3\xi + \rho)\vartheta \leq 1/10. \tag{8.6}$$

Then for any $K \geq K_0$ with K_0 depending only on ϑ and Φ the following statement holds. For any Z and ℓ as above if

$$\Big[\frac{1}{M^2}\int_{-2M}^0 \mathrm{d}t \sum_i (u_i(t) - \psi_i)_+^2\Big]^{1/2} \leq \varepsilon_0 \ell, \qquad \psi_i = \psi_i^{(M,Z,\ell)}, \tag{8.7}$$

(ε_0 was given in the beginning of the this theorem) then for any $|i| \leq K$ we have

$$\sup_{t\in[-M,0]} u_i(t) \leq \frac{\ell}{2} + \psi_i \tag{8.8}$$

We remark that the constant ε_0 is independent of any other parameter, in particular it is independent of $\ell, Z, M, c_1, \xi, \rho, \vartheta$.

The bound (8.8) is essentially an $L^2 \to L^\infty$ estimate, which is the basic input for applying the method in [10]. Notice that (8.2), (8.3) and (8.4) are the same as (7.59) (7.60) and (7.79) but ϱ_1 and ξ' are replaced by ϱ and ξ for simplicity of notations.

Proof. We first prove a weaker version of this result, namely, the following lemma:

Lemma 8.2 *Under the assumptions and notations of Theorem 8.1 we have*

$$\sup_{t\in[-M,0]} \sum_i (u_i(t) - \psi_i^{\ell/3})_+^2 \leq CM^\Phi \ell^2. \tag{8.9}$$

where Φ is defined in (8.6).

Proof. Assume for notational simplicity that $Z = 0$ (this requires relabelling the configuration space). By direct computation, we have

$$\partial_t \frac{1}{2} \sum_i [u_i - \psi_i^\ell]_+^2 = -\sum_{ij} (u_i - \psi_i^\ell)_+ B_{ij}(u_i - u_j) - \sum_i (u_i - \psi_i^\ell)_+ W_i u_i \tag{8.10}$$

Recall that B_{ij} depends on time, but we will omit this from the notation. Since $W_i \geq 0$, the last term can be bounded by

$$-\sum_i (u_i - \psi_i^\ell)_+ W_i u_i \leq -\sum_i (u_i - \psi_i^\ell)_+ W_i (u_i - \psi_i^\ell)_+ = -\mathfrak{w}[(u - \psi^\ell)_+, (u - \psi^\ell)_+].$$

In the first term on the right hand side of (8.10) we can symmetrize and then add and subtract ψ^ℓ to u we get

$$\begin{aligned} -\sum_{ij} (u_i - \psi_i^\ell)_+ B_{ij}(u_i - u_j) &= -\,\mathfrak{b}[(u - \psi^\ell)_+, u] \\ &= -\,\mathfrak{b}[(u - \psi^\ell)_+, (u - \psi^\ell)_+] - \mathfrak{b}[(u - \psi^\ell)_+, (u - \psi^\ell)_-] - \mathfrak{b}[(u - \psi^\ell)_+, \psi^\ell] \end{aligned}$$

Since $B_{ij} \geq 0$ and $[a_+ - b_+][a_- - b_-] \geq 0$ for any real numbers a, b, for the cross-term we have $\mathfrak{b}[(u - \psi^\ell)_+, (u - \psi^\ell)_-] \geq 0$. Thus the last equation is bounded by

$$\leq -\mathfrak{b}[(u - \psi^\ell)_+, (u - \psi^\ell)_+] - \mathfrak{b}[(u - \psi^\ell)_+, \psi^\ell]. \tag{8.11}$$

Using the definition of $\mathfrak{a}$ (7.16), we have thus proved that

$$\partial_t \frac{1}{2} \sum_i [u_i - \psi_i^\ell]_+^2 \leq -\mathfrak{a}[(u - \psi^\ell)_+, (u - \psi^\ell)_+] - \mathfrak{b}[(u - \psi^\ell)_+, \psi^\ell].$$

Decompose the error term into

$$\mathfrak{b}[(u - \psi^\ell)_+, \psi^\ell] = \Omega_1 + \Omega_2 + \Omega_3,$$

$$\Omega_1 := \frac{1}{2} \sum_{|i-j| \geq M} B_{ij} [\psi_i^\ell - \psi_j^\ell] \big((u_i - \psi_i^\ell)_+ - (u_j - \psi_j^\ell)_+ \big)$$

and Ω_2 and Ω_3 are defined in the same way except that the summation is restricted to $\widehat{C} K^\xi \leq |i - j| \leq M$ for Ω_2 and $|i - j| \leq \widehat{C} K^\xi$ for Ω_3, where $\widehat{C}$ is the constant from (7.78) and (7.79). We have

$$|\psi_i^\ell - \psi_j^\ell| \leq \frac{\ell |i - j|}{\sqrt{M} \left[\sqrt{|i|} + \sqrt{|j|} \right]} \tag{8.12}$$

The last term is bounded by $\ell M^{-1/2} |i - j|^{1/2}$. Recall from (7.79) that in the regime $|i - j| \geq M$ we have $B_{ij} \leq \frac{C}{|i-j|^2}$ Together with $M \geq K^\xi$, we have

$$|\Omega_1| \leq \ell M^{-1/2} \sum_{|i-j| \geq M} \frac{1}{|i - j|^{3/2}} \left[(u_i - \psi_i^\ell)_+ + (u_j - \psi_j^\ell)_+ \right] \leq \frac{\ell}{M} \sum_i (u_i - \psi_i^\ell)_+$$

For Ω_2, by symmetry of B_{ij} we can rewrite it as

$$\begin{aligned}
-\Omega_2 := &- \sum_{\widehat{C}K^\xi \le |i-j| \le M, \psi_i^\ell \le \psi_j^\ell} B_{ij}[\psi_i^\ell - \psi_j^\ell]\big((u_i - \psi_i^\ell)_+ - (u_j - \psi_j^\ell)_+\big) \\
\le &- \sum_{\widehat{C}K^\xi \le |i-j| \le M, \psi_i^\ell \le \psi_j^\ell} B_{ij}[\psi_i^\ell - \psi_j^\ell]\,\big[(u_i - \psi_i^\ell)_+ - (u_j - \psi_j^\ell)_+\big]\chi(u_i - \psi_i^\ell > 0) \\
\le &\frac{1}{4} \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}\big[(u_i - \psi_i^\ell)_+ - (u_j - \psi_j^\ell)_+\big]^2 \\
&+ 4 \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}|\psi_i^\ell - \psi_j^\ell|^2\chi(u_i - \psi_i^\ell > 0).
\end{aligned}$$

The first term is $\frac{1}{2}\mathfrak{b}[(u-\psi^\ell)_+, (u-\psi^\ell)_+]$ and can be absorbed in the first term on the r.h.s. of (8.11). By (8.12) and (7.79), the second term is bounded by

$$4 \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}|\psi_i^\ell - \psi_j^\ell|^2\chi(u_i - \psi_i^\ell > 0) \le C\ell^2 M^{-1} \sum_i \chi(u_i - \psi_i^\ell > 0), \tag{8.13}$$

i.e.

$$\begin{aligned}
-\Omega_2 \le &\frac{1}{2}\mathfrak{b}[(u-\psi^\ell)_+, (u-\psi^\ell)_+] + C\ell^2 M^{-1} \sum_i \chi(u_i - \psi_i^\ell > 0) \\
\le &\frac{1}{2}\mathfrak{a}[(u-\psi^\ell)_+, (u-\psi^\ell)_+] + C\ell^2 M^{-1} \sum_i \chi(u_i - \psi_i^\ell > 0)
\end{aligned}$$

using that $\mathfrak{b} \le \mathfrak{a}$.

A similar estimate is performed for Ω_3, but in the corresponding last term we use that

$$|\psi_i^\ell - \psi_j^\ell| \le C(\ell/M)^2$$

for $|i-j| \le K^\xi$. Thus we have

$$\begin{aligned}
-\Omega_3 \le &\sum_{|i-j| \le \widehat{C}K^\xi} B_{ij}|\psi_i^\ell - \psi_j^\ell|^2\chi(u_i - \psi_i^\ell > 0) \\
\le &CK^{2\xi}(\ell/M)^2 \sum_{|i-j| \le \widehat{C}K^\xi} \chi(u_i - \psi_i^\ell > 0)B_{ij} \\
\le &C\frac{K^{3\xi}\ell^2}{M^2} \sum_i \chi(u_i - \psi_i^\ell > 0)[B_{i,i+1} + B_{i,i-1}]
\end{aligned} \tag{8.14}$$

Here we just overestimated sums by $\widehat{C}K^\xi$. The conclusion of the energy estimate is

$$\begin{aligned}
\partial_t \frac{1}{2} \sum_i [u_i - \psi_i^\ell]_+^2 \le &- \frac{1}{2}\mathfrak{a}[(u-\psi^\ell)_+, (u-\psi^\ell)_+] \\
&+ \frac{C\ell}{M} \sum_i (u_i - \psi_i^\ell)_+ + \frac{C\ell^2}{M} \sum_i \chi(u_i - \psi_i^\ell > 0) + \Omega_4
\end{aligned} \tag{8.15}$$

$$\Omega_4 = \frac{CK^{3\xi}\ell^2}{M^2}\sum_i \chi(u_i - \psi_i^\ell > 0)[B_{i,i+1} + B_{i,i-1}] \tag{8.16}$$

Due to the assumption (8.5), we can assume that the summation of i is restricted to $|i| \le M^{1+3c_1}$.

By (8.2), we have

$$\int_\tau^t \Omega_4 \mathrm{d}s \le \frac{CK^{3\xi}\ell^2}{M^{1-3c_1}}\int_\tau^t \mathrm{d}s \frac{1}{M^{1+3c_1}}\sum_{|i|\le M^{1+3c_1}}[B_{i,i+1}+B_{i,i-1}](s) \le (t-\tau)K^{3\xi+\rho}\ell^2 M^{3c_1-1}. \tag{8.17}$$

For a fixed M, let $T_k = -M(1+2^{-k})$, $\ell_k = \frac{\ell}{3}(1-2^{-k}) \nearrow \frac{\ell}{3}$, $Q_k = [T_k, 0]\times\{1,2,\dots K\}$. Define

$$U_k = \sup_{t\in[T_k,0]} \frac{1}{M\ell_k^2}\sum_i (u_i - \psi_i^{\ell_k})_+^2(t) + \frac{1}{M\ell_k^2}\int_{T_k}^0 \mathfrak{a}[(u-\psi^{\ell_k})_+, (u-\psi^{\ell_k})_+](s)\mathrm{d}s.$$

The prefactors $1/(M\ell_k^2)$ just make the size of U order one.

Integrating (8.15) from τ to t with $T_{k-1}\le\tau\le T_k \le t\le 0$, we have from (8.17)

$$\begin{aligned}
&\sum_i [u_i - \psi_i^{\ell_k}]_+^2(t) + \int_\tau^t \mathfrak{a}[(u-\psi^{\ell_k})_+, (u-\psi^{\ell_k})_+](s)\mathrm{d}s \\
&\qquad \le \sum_i [u_i - \psi_i^{\ell_k}]_+^2(\tau) + C\int_\tau^t \left[\frac{\ell_k}{M}\sum_i (u_i-\psi_i^{\ell_k})_+(s) + \frac{\ell_k^2}{M}\sum_i \chi(u_i - \psi_i^{\ell_k} > 0)(s)\right]\mathrm{d}s \\
&\qquad\quad + C(t-\tau)K^{3\xi+\rho}\ell^2 M^{3c_1-1}.
\end{aligned} \tag{8.18}$$

Taking the average over $\tau\in[T_{k-1}, T_{k-1}+2^{-k-1}M] = [T_{k-1}, T_k - 2^{-k-1}M]$ and using that in this regime $\tau - t \sim 2^{-k-1}M$, we have

$$\begin{aligned}
&\sum_i [u_i - \psi_i^{\ell_k}]_+^2(t) + \int_{T_k}^t \mathfrak{a}[(u-\psi^{\ell_k})_+, (u-\psi^{\ell_k})_+](s)\mathrm{d}s \\
&\le C\frac{2^{k+1}}{M}\int_{T_{k-1}}^{T_k - 2^{-k-1}M}\sum_i [u_i-\psi_i^{\ell_k}]_+^2(s)\mathrm{d}s \\
&\quad + C\int_{T_{k-1}}^t \left[\frac{\ell_k}{M}\sum_i (u_i-\psi_i^{\ell_k})_+(s) + \frac{\ell_k^2}{M}\sum_i \chi(u_i - \psi_i^{\ell_k} > 0)(s)\right]\mathrm{d}s + CK^{3\xi+\rho}\ell^2 M^{3c_1}
\end{aligned}$$

Dividing through by $M\ell_k^2$ and taking supremum over $t\in[T_k, 0]$, for $k\ge 1$ we have

$$\begin{aligned}
U_k \le & C\frac{2^{k+1}}{M^2}\int_{T_{k-1}}^0 \sum_i \Big[\frac{1}{\ell_k^2}[u_i-\psi_i^{\ell_k}]_+^2 + \frac{1}{\ell_k}(u_i-\psi_i^{\ell_k})_+ + \chi(u_i-\psi_i^{\ell_k}>0)\Big](s)\mathrm{d}s \\
& + M^{3c_1}\frac{CK^{3\xi+\rho}}{M}
\end{aligned} \tag{8.19}$$

The first three integrands have the same scaling dimensions as u^2/ℓ^2. One key idea is to estimate these terms in terms of the L^4 norm of u and then using the Sobolev inequality. It is elementary to check these three integrands can be bounded by the L^4 norm of $(u-\psi^{\ell_k})_+$, by using

that if $u_i \geq \psi_i^{\ell_k}$, then $u_i - \psi_i^{\ell_{k-1}} \geq \ell_k - \ell_{k-1} = 2^{-k}\frac{\ell}{3} \geq 2^{-(k+2)}\ell$:

$$
\begin{aligned}
\sum_i (u_i - \psi_i^{\ell_k})_+ \leq & \sum_i (u_i - \psi_i^{\ell_k})_+ \chi(u_i - \psi_i^{\ell_{k-1}} > 2^{-(k+2)}\ell) \\
\leq & (2^{k+1})^3 \ell_k^{-3} \sum_i (u_i - \psi_i^{\ell_{k-1}})_+^4, \\
\sum_i \chi(u_i - \psi_i^{\ell_k} > 0) \leq & (2^{k+2})^4 \ell_k^{-4} \sum_i (u_i - \psi_i^{\ell_{k-1}})_+^4, \\
\sum_i [u_i - \psi_i^{\ell_k}]_+^2 \leq & (2^{k+2})^2 \ell_k^{-2} \sum_i (u_i - \psi_i^{\ell_{k-1}})_+^4.
\end{aligned}
\tag{8.20}
$$

Recall the Sobolev inequality (7.45) $\Big(\sum_i f_i^4\Big)^{1/2} \leq \Big(\sum_i f_i^2\Big)^{1/2} \mathfrak{a}[f,f]^{1/2}$. Thus we have proved that for $k \geq 2$

$$
\begin{aligned}
U_k \leq & (2^{k+2})^5 \frac{1}{M^2 \ell_k^4} \int_{T_{k-1}}^0 \mathrm{d}s \Big[\sum_i (u_i - \psi_i^{\ell_{k-1}})_+^2(s) \Big] \mathfrak{a}[(u - \psi^{\ell_{k-1}})_+, (u - \psi^{\ell_{k-1}})_+](s) \\
& + C M^{3c_1 - 1} K^{3\xi + \rho} \\
\leq & C_1^k U_{k-1}^2 + M^{-1+\Phi}, \quad k \geq 2.
\end{aligned}
\tag{8.21}
$$

recalling that $K = M^\vartheta$ and $\Phi = 3c_1 + (3\xi + \rho)\vartheta$.

For $k = 1$, we estimate the integrands in (8.19) by ℓ^2-norms. We have the following general estimates for any $\ell' < \ell''$

$$
\begin{aligned}
\sum_i (u_i - \psi_i^{\ell''})_+ \leq \sum_i (u_i - \psi_i^{\ell'})_+ \chi(u_i - \psi_i^{\ell'} > \ell'' - \ell') \leq \frac{1}{\ell'' - \ell'} \sum_i (u_i - \psi_i^{\ell'})_+^2 \\
\sum_i \chi(u_i - \psi_i^{\ell''} > 0) \leq \frac{1}{(\ell'' - \ell')^2} \sum_i (u_i - \psi_i^{\ell'})_+^2.
\end{aligned}
\tag{8.22}
$$

We use (8.22) with $\ell'' = \ell_1$ and $\ell' = 0$ in (8.19), this implies that

$$
U_1 \leq \frac{C}{\ell_1^2 M^2} \int_{-2M}^0 \sum_i (u_i - \psi_i)_+^2 + M^{-1+\Phi}.
$$

Without loss of generality, we assume that $C_1 \geq 2$, where C_1 is the constant in (8.21). Now choose the universal constant ε_0 in (8.7) so small and M big enough so that this last inequality implies

$$
U_1 \leq \frac{1}{100 C_1^6}.
$$

One can check the recurrence relation (8.21) implies that

$$
U_k \leq C M^{-1+\Phi} \tag{8.23}
$$

for any k such that $C_1^{-k} \leq M^{-1+\Phi}$. Choosing k sufficiently large, noticing that U_k is decreasing in k (as T_k and ℓ_k increase), we find from the ℓ^2-norm term in U_k that (8.9) holds. This proves Lemma 8.2. □

We now repeat the proof of Lemma 8.2 but with $\ell_k, k \geq 1$, replaced by

$$\widehat{\ell}_k = \frac{2\ell}{5}(1 - 2^{-k-2}) \tag{8.24}$$

and with integration on a smaller time interval. Choose M_1 such that $M^{\Phi} \ll M_1 \ll M^{1-\Phi}$, and set $\widehat{T}_k := -M_1(1 + 2^{-k})$. Define

$$\widehat{U}_k = \sup_{t \in [\widehat{T}_k, 0]} \frac{1}{M\widehat{\ell}_k^2} \sum_i (u_i - \psi_i^{\widehat{\ell}_k})_+^2(t) + \frac{1}{M\widehat{\ell}_k^2} \int_{\widehat{T}_k}^0 \mathfrak{a}[(u - \psi^{\widehat{\ell}_k})_+, (u - \psi^{\widehat{\ell}_k})_+](s)\mathrm{d}s.$$

The proof is unchanged up to (8.17), the integral of

$$\widehat{\Omega}_4(s) := \frac{CK^{3\xi}\widehat{\ell}^2}{M^2} \sum_i \chi(u_i(s) - \psi_i^{\widehat{\ell}} > 0)[B_{i,i+1}(s) + B_{i,i-1}(s)].$$

is still estimated by

$$\int_\tau^t \widehat{\Omega}_4(s)\mathrm{d}s \leq C(t - \tau)K^{3\xi+\rho}\widehat{\ell}^2 M^{-1+3c_1} \leq C(t - \tau)\widehat{\ell}^2 M^{-1+\Phi}$$

Similarly to (8.18), we integrate (8.15) (with $\widehat{\ell}$ replacing ℓ) from τ to t with $\widehat{T}_{k-1} \leq \tau \leq \widehat{T}_k \leq t \leq 0$

$$\begin{aligned}
&\sum_i [u_i - \psi_i^{\widehat{\ell}_k}]_+^2(t) + \int_\tau^t \mathfrak{a}[(u - \psi^{\widehat{\ell}_k})_+, (u - \psi^{\widehat{\ell}_k})_+](s)\mathrm{d}s \\
&\qquad \leq \sum_i [u_i - \psi_i^{\widehat{\ell}_k}]_+^2(\tau) + C\int_\tau^t \left[\frac{\ell_k}{M}\sum_i (u_i - \psi_i^{\widehat{\ell}_k})_+(s) + \frac{\widehat{\ell}_k^2}{M}\sum_i \chi(u_i - \psi_i^{\widehat{\ell}_k} > 0)(s)\right]\mathrm{d}s \\
&\qquad\quad + C(t - \tau)\ell^2 M^{-1+\Phi}.
\end{aligned} \tag{8.25}$$

Taking the average over $\tau \in [\widehat{T}_{k-1}, \widehat{T}_{k-1} + 2^{-k-1}M_1] = [\widehat{T}_{k-1}, \widehat{T}_k - 2^{-k-1}M_1]$ and using that in this regime $\tau - t \sim 2^{-k-1}M_1$, we have

$$\begin{aligned}
&\sum_i [u_i - \psi_i^{\widehat{\ell}_k}]_+^2(t) + \int_{\widehat{T}_k}^t \mathfrak{a}[(u - \psi^{\widehat{\ell}_k})_+, (u - \psi^{\widehat{\ell}_k})_+](s)\mathrm{d}s \\
&\leq C\frac{2^{k+1}}{M_1} \int_{\widehat{T}_{k-1}}^{\widehat{T}_k - 2^{-k-1}M_1} \sum_i [u_i - \psi_i^{\widehat{\ell}_k}]_+^2(s)\mathrm{d}s \\
&\quad + C\int_{\widehat{T}_{k-1}}^t \left[\frac{\widehat{\ell}_k}{M}\sum_i (u_i - \psi_i^{\widehat{\ell}_k})_+(s) + \frac{\widehat{\ell}_k^2}{M}\sum_i \chi(u_i - \psi_i^{\widehat{\ell}_k} > 0)(s)\right]\mathrm{d}s + C\ell^2 M_1 M^{-1+\Phi}
\end{aligned}$$

Dividing through by $M\widehat{\ell}_k^2$ and taking supremum over $t \in [\widehat{T}_k, 0]$, for $k \geq 1$ and using $M_1 \leq M$, we have

$$\begin{aligned}
\widehat{U}_k \leq & C\frac{2^{k+1}}{MM_1} \int_{\widehat{T}_{k-1}}^0 \sum_i \Big[\frac{1}{\widehat{\ell}_k^2}[u_i - \psi_i^{\widehat{\ell}_k}]_+^2 + \frac{1}{\widehat{\ell}_k}(u_i - \psi_i^{\widehat{\ell}_k})_+ + \chi(u_i - \psi_i^{\widehat{\ell}_k} > 0)\Big](s)\mathrm{d}s \\
& + CM_1 M^{-2+\Phi}
\end{aligned} \tag{8.26}$$

Using the bounds (8.20) and the Sobolev inequality, we get, instead of (8.21) for any $k \geq 2$

$$\begin{aligned}\widehat{U}_k \leq & \frac{(2^{k+2})^5}{MM_1\widehat{\ell}_k^4}\int_{\widehat{T}_{k-1}}^0 \mathrm{d}s\Big[\sum_i (u_i-\psi_i^{\widehat{\ell}_{k-1}})_+^2(s)\Big]\mathfrak{a}[(u-\psi^{\widehat{\ell}_{k-1}})_+,(u-\psi^{\widehat{\ell}_{k-1}})_+](s)\\ & + CM_1M^{-2+\Phi}\\ \leq & C_1^k\frac{M}{M_1}\widehat{U}_{k-1}^2 + CM_1M^{-2+\Phi}, \quad k\geq 2.\end{aligned} \tag{8.27}$$

This new recurrence inequality has the solution

$$\widehat{U}_{k_0} \leq CM_1M^{-2+\Phi} \tag{8.28}$$

for a sufficiently large k_0, as long as the recursion can be started, i.e. if we knew

$$\widehat{U}_1 \ll \frac{M_1}{M}. \tag{8.29}$$

For $k=1$ the estimate (8.26) together with (8.22) (with $\widehat{\ell}_1$ replacing ℓ_1) becomes

$$\begin{aligned}\widehat{U}_1 \leq & \frac{C}{MM_1}\int_{-2M_1}^0 \sum_i \Big[\frac{1}{\widehat{\ell}_1^2}[u_i-\psi_i^{\widehat{\ell}_1}]_+^2 + \frac{1}{\widehat{\ell}_1}(u_i-\psi_i^{\widehat{\ell}_1})_+ + \chi(u_i-\psi_i^{\widehat{\ell}_1}>0)\Big](s)\mathrm{d}s + CM_1M^{-2+a}\\ \leq & \frac{C}{MM_1}\int_{-2M_1}^0 \sum_i \frac{1}{\ell^2}[u_i(s)-\psi_i^{\ell/3}]_+^2\mathrm{d}s + CM_1M^{-2+\Phi}\\ \leq & \frac{CM^a}{M} + CM_1M^{-2+\Phi}\end{aligned}$$

In the second step we used (8.22) with $\ell'' = \widehat{\ell}_1$ and $\ell' = \ell/3$ noting that $\widehat{\ell}_1 = \frac{7}{20}\ell > \frac{1}{3}\ell$. In the last step we used (8.9) and $2M_1 \leq M$. Thus (8.29) is satisfied if $M_1 \gg M^\Phi$.

Finally, (8.28) implies

$$\sup_{t\in[-M_1,0]}\sum_i (u_i(t)-\psi_i^{2\ell/5})_+^2 \leq C\ell^2M_1M^{-1+\Phi} = o(\ell^2)$$

since $M_1 \ll M^{1-\Phi}$, i.e.

$$\sup_{t\in[-M_1,0]}\sup_i (u_i(t)-\psi_i^{2\ell/5})_+ \leq o(\ell)$$

The estimate (8.8) now follows if we repeat the same argument for each time interval of length M_1 within $[-M,0]$. This completes the proof of Theorem 8.1. $\square$

For any Z, M and ℓ define

$$\bar{\psi}_i = \bar{\psi}_i^{(M,Z,\ell)} := \ell\Big(\Big|\frac{i-Z}{M}\Big|^{1/4} - 1\Big)_+. \tag{8.30}$$

We have the following corollary

Corollary 8.3 *There exists a constant $\delta > 0$ (depending only on ε_0 from Theorem 8.1) such that for any $K \geq K_0(\vartheta, \Phi)$ if for some M, Z and ℓ a solution u satisfies* (8.5),

$$\sup_{t\in[-2M,0]} \max_i u_i(t) \leq \ell + \bar{\psi}_i \tag{8.31}$$

and

$$\frac{1}{M^2}\int_{-2M}^{0} \#\Big\{i \in [Z-2M, Z+2M] \ : \ u_i(t) > 0\Big\}\mathrm{d}t \leq \delta, \tag{8.32}$$

then

$$\sup_{t\in[-M,0]} \max_{i\in[\![Z-M,Z+M]\!]} u_i(t) \leq \frac{\ell}{2}. \tag{8.33}$$

Proof. Fix $(t_0, i_0) \in [-M, 0] \times [\![-M, M]\!]$, our goal is to show that $u_{i_0}(t_0) \leq \ell/2$. Consider the interval $I_0 = [\![x_0 - M/R, x_0 + M/R]\!]$ of length $2M/R$ about x_0, where R is a large constant. We will apply Theorem 8.1 to the interval I_0, i.e. we define

$$\psi_i := \psi_i^{(M', i_0, \ell)} = \ell\Big(\Big|\frac{i-i_0}{M'}\Big|^{1/2} - 1\Big)_+, \qquad M' := M/R.$$

Since $\psi_{i_0} = 0$, the estimate $u_{i_0}(t_0) \leq \ell/2$ would follow from the conclusion of Theorem 8.1. We just have to check that

$$\Big[\frac{1}{(M')^2}\int_{-2M'}^{0}\sum_i (u_i - \psi_i)_+^2\Big]^{1/2} \leq \varepsilon_0 \ell \tag{8.34}$$

if δ is sufficiently small (depending only on ε_0). Notice that if R is a sufficiently large universal constant, then

$$\ell + \bar{\psi}_i \leq \psi_i, \qquad \text{for } \ |i| \geq 2M. \tag{8.35}$$

This follows from the simple inequality

$$1 + \Big(\Big|\frac{i}{M}\Big|^{1/4} - 1\Big)_+ \leq \Big(\Big|\frac{i-i_0}{M/R}\Big|^{1/2} - 1\Big)_+$$

whenever $|i_0| \leq M$ and $|i| \geq 2M$. Thus we have

$$\int_{-2M'}^{0}\sum_i (u_i(t) - \psi_i)_+^2 \mathrm{d}t = \int_{-2M'}^{0}\sum_{|i|\leq 2M} (u_i(t) - \psi_i)_+^2 \mathrm{d}t \leq M^2\delta \max_{|i|\leq 2M} (\ell + \bar{\psi}_i)^2 \leq C_2 M^2 \ell^2 \delta.$$

with some universal C_2. In the first equality we used that for $|i| \geq 2M$ the condition (8.31) and (8.35) imply that $u_i(t) \leq \psi_i$. In the second inequality we used that $(u_i - \psi_i)_+ \leq (u_i)_+ \leq \ell + \bar{\psi}_i$ and (8.32). The last inequality is a simple arithmetic estimate. Thus (8.34) follows if δ is so small that

$$C_2 R^2 \delta \leq \varepsilon_0^2.$$

This completes the proof of the Corollary. □

Remark 8.4 *Suppose that in addition to* (8.3), (8.2) *and* (8.4), *the bound*

$$W_i \leq \frac{K^\xi}{d_i}, \quad \textit{if} \quad d_i \geq K^{C\xi} \tag{8.36}$$

analogous to (7.80) *is also satisfied. For any constant* $|A| \le \ell M^{1+c_1}$*, let* $v = u - A$*. Then* v *satisfies*

$$\partial_s v_i(s) = -(\mathcal{A}(s)v(s))_i - AW_i \tag{8.37}$$

We can apply the proof for Theorem 8.1 to v *and view* $W_i A$ *as an error term. Then instead of* (8.8)*, we will have*

$$\sup_{t\in[-M,0]} v_i(t) \le \frac{\ell}{2} + \psi_i + \ell M^{1+c_1} K^{-1+C\xi} \le \frac{\ell}{2} + \psi_i + \ell K^{-1/2} \tag{8.38}$$

provided that $\vartheta \ge 3$*. The last error term* $\ell K^{-1/2}$ *is much smaller than the Hölder continuity bound we aim for Theorem 7.7 and can be neglected. Similar remark applies to Corollary 8.3.*

8.2 Local decrease of oscillation: Second De Giorgi lemma

As before, we are given three parameters, M, Z, ℓ. The center Z can be set zero for notational simplicity (this requires relabelling the configuration space). Define a nonpositive function supported on $|i| \le 3M$ by

$$F_i = F_i^{(M,Z,\ell)} = \ell \cdot \max\Big\{ -1, \min\Big(0, \Big|\frac{i-Z}{M}\Big|^2 - 9\Big)\Big\}.$$

Notice that $-\ell \le F \le 0$ and $F_i = 0$ if $|i - Z| \ge 3M$. For any $\lambda > 0$ small, set

$$\widetilde{\psi}_i = \widetilde{\psi}_i^{(M,Z,\ell,\lambda)} := \ell\Big[\Big(\Big|\frac{i-Z}{M}\Big| - \lambda^{-4}\Big)^{1/4} - 1\Big]_+ \qquad \text{for} \qquad |i - Z| \ge M\lambda^{-4} \tag{8.39}$$

and $\widetilde{\psi}_i = 0$ otherwise. We also define three cutoffs, all depending on all four parameters, M, Z, ℓ, λ:

$$\varphi_i^{(0)} = \ell + \widetilde{\psi}_i + F_i$$

$$\varphi_i^{(1)} = \ell + \widetilde{\psi}_i + \lambda F_i$$

$$\varphi_i^{(2)} = \ell + \widetilde{\psi}_i + \lambda^2 F_i$$

Notice that

$$\varphi_i^{(0)} \le \varphi_i^{(1)} \le \varphi_i^{(2)} \le \ell + \widetilde{\psi}_i, \tag{8.40}$$

and when $|i| \ge 3M$ all inequalities becomes equalities. Notice that $\varphi_i^{(k)} = 0$ if $|i - Z| \le 2\sqrt{2}M$ for $k = 0, 1, 2$.

Lemma 8.5 *Let* $M^\vartheta = K$ *and assume that* (8.2)*,* (8.3) *and* (8.4) *hold with some* ρ *and* ξ *so that* (8.6) *holds. Recall the constant* δ *from Corollary 8.3 which depends only on* ε_0 *from Theorem 8.1. Depending only on this* δ*, there exists* $\mu > 0$*,* $\gamma > 0$ *and* $\lambda \in (0, 1/8)$*, such that whenever a solution* $u(t)$ *to* (7.57) *satisfies*

$$u_i(t) \le \ell + \widetilde{\psi}_i, \qquad t \in [-3M, 0], \quad \forall i \tag{8.41}$$

$$\frac{1}{M^2}\int_{-3M}^{-2M} \#\Big\{i \in [Z - M, Z + M] \,:\, u_i(t) < \varphi_i^{(0)}\Big\}\mathrm{d}t \ge \mu, \tag{8.42}$$

and

$$\frac{1}{M^2}\int_{-2M}^{0} \#\Big\{i \,:\, u_i(t) > \varphi_i^{(2)}\Big\}\mathrm{d}t \ge \delta, \tag{8.43}$$

then

$$\frac{1}{M^2}\int_{-3M}^{0}\#\Big\{i\ :\ \varphi_i^{(0)}<u_i(t)<\varphi_i^{(2)}\Big\}\mathrm{d}t\geq\gamma. \tag{8.44}$$

If in addition (8.36) *is satisfied and* v *is a solution* (8.37) *satisfying* (8.41), (8.42) *and* (8.43), *then* (8.44) *holds.*

This Lemma asserts that whenever the function u increases from $\varphi^{(0)}$ to $\varphi^{(2)}$ in time of order M, then there is a time interval of order M so that u is between $\varphi^{(0)}$ and $\varphi^{(2)}$. The rest of this section is devoted to the proof of this Lemma.

Proof. We can fix $\mu<1/8$. The parameter δ may have to be reduced a bit along the proof. We are looking for a sufficiently small λ so that there will be a positive γ. The key ingredient of the proof is an energy inequality (8.48) including a new dissipation term which was dropped in the previous section. We first derive this inequality.

8.2.1 Dissipation with the good term

Let $-3M\leq T_1<T_2<0$. We use (8.11) (with cutoff $\varphi^{(1)}$) but we keep the good term

$$\frac{1}{2}\sum_i[u_i(t)-\varphi_i^{(1)}]_+^2\Big|_{t=T_1}^{T_2}+\int_{T_1}^{T_2}\mathfrak{b}[(u(t)-\varphi^{(1)})_+,(u(t)-\varphi^{(1)})_+]\mathrm{d}t \tag{8.45}$$

$$=-\int_{T_1}^{T_2}\mathfrak{b}[(u(t)-\varphi^{(1)})_+,(u(t)-\varphi^{(1)})_-]\mathrm{d}t-\int_{T_1}^{T_2}\mathfrak{b}[(u(t)-\varphi^{(1)})_+\theta,\varphi^{(1)}]\mathrm{d}t$$

Notice that we inserted a characteristic function $\theta_i=\mathbf{1}(|i|\leq 3M)$ using the fact that (8.41) and (8.40) imply $u_i\leq\varphi_i^{(1)}$ for $|i|\geq 3M$, i.e. $u_i-\varphi^{(1)}=(u_i-\varphi^{(1)})\theta$. The last error term will be Schwarzed and partly absorbed in the quadratic term in the left. Here we use

$$\mathfrak{b}(f\theta,g)=\sum_{ij}(f_i\theta_i-f_j\theta_j)B_{ij}(g_i-g_j)=\sum_{ij}(f_i\theta_i-f_j\theta_j)(\theta_i+\theta_j-\theta_i\theta_j)B_{ij}(g_i-g_j)$$

so

$$|\mathfrak{b}(f\theta,g)|\leq\sum_{ij}(f_i\theta_i-f_j\theta_j)^2B_{ij}+\sum_{ij}\theta_iB_{ij}(g_i-g_j)^2$$

In the other term we use again a Schwarz to separate the two summands in $\varphi^{(1)}$. We are left with to control

$$\int_{T_1}^{T_2}\left[\lambda^2\sum_{i,j}(F_i-F_j)^2B_{ij}+\sum_{i,j}(\widetilde{\psi}_i-\widetilde{\psi}_j)^2B_{ij}\theta_i\right](t)\mathrm{d}t \tag{8.46}$$

Since $|F_i-F_j|\leq C\ell M^{-1}|i-j|$ and supported on $|i|,|j|\leq 3M$, by splitting the summation to $|i-j|\leq K^\xi$ regime and its complement, we can bound the first term by

$$\int_{T_1}^{T_2}\lambda^2\sum_{i,j}(F_i-F_j)^2B_{ij}(t)\mathrm{d}t\leq\lambda^2\ell^2M^{-2}\int_{T_1}^{T_2}\sum_{|i|,|j|\leq 3M}|i-j|^2B_{ij}(t)$$

$$\leq\lambda^2\ell^2M^{-2}K^{3\xi}\int_{T_1}^{T_2}\sum_{|i|\leq 3M}B_{i,i+1}(t)+C\lambda^2\ell^2M^{-2}\int_{T_1}^{T_2}\sum_{\substack{|i|,|j|\leq 3M\\ |i-j|\geq K^\xi}}\frac{|i-j|^2}{|i-j|^2}$$

where we have used $B_{i,j} \le B_{i,i+1}$ in the first regime and (8.4) in the other regime. By (8.2), we can bound the last line by

$$\lambda^2\ell^2 K^{3\xi+\rho} + \lambda^2\ell^2(T_2 - T_1) \le C\lambda^2\ell^2(T_2 - T_1).$$

For the second term in (8.46), we use that $\widetilde{\psi}_i\theta_i = 0$ and the supports of θ_i and $\widetilde{\psi}_j$ are separated by a distance of order $M \gg K^\xi$. Thus we can use (8.4) to estimate the kernel:

$$\int_{T_1}^{T_2} \sum_{i,j} (\widetilde{\psi}_i - \widetilde{\psi}_j)^2 B_{ij}(t)\theta_i \mathrm{d}t \le C \int_{T_1}^{T_2} \sum_{|i|\le 3M} \sum_{|j|\ge M\lambda^{-4}} \frac{\widetilde{\psi}_j^2}{|i-j|^2}\mathrm{d}t$$

$$\le M(T_2 - T_1) \sum_{|j|\ge M\lambda^{-4}} \frac{\widetilde{\psi}_j^2}{|j|^2} \le \ell^2\lambda^2(T_2 - T_1),$$

where we have used $\widetilde{\psi}_j \sim \ell(j/M)^{1/4}$ for large j.

Inserting these two error estimates into (8.45), we have

$$\frac{1}{2}\sum_i [u_i(t) - \varphi_i^{(1)}]_+^2 \Big|_{t=T_1}^{T_2} + \frac{1}{2}\int_{T_1}^{T_2} \mathfrak{b}[(u(t) - \varphi^{(1)})_+, (u(t) - \varphi^{(1)})_+]\mathrm{d}t \tag{8.47}$$

$$\le -\int_{T_1}^{T_2} \mathfrak{b}[(u(t) - \varphi^{(1)})_+, (u(t) - \varphi^{(1)})_-]\mathrm{d}t + C\ell^2\lambda^2(T_2 - T_1).$$

Define

$$H(t) = \sum_i (u_i(t) - \varphi_i^{(1)})_+^2.$$

Since $|u_i(t) - \varphi_i^{(1)}| \le \lambda\theta_i$, we have, for all t,

$$H(t) \le C\lambda^2 M.$$

Hence we have

$$H(T_2) + \int_{T_1}^{T_2} \mathfrak{b}[(u(t) - \varphi^{(1)})_+, (u(t) - \varphi^{(1)})_-]\mathrm{d}t \le H(T_1) + C\ell^2\lambda^2(T_2 - T_1) \le C\ell^2\lambda^2 M \tag{8.48}$$

for any $-3M \le T_1 < T_2 < 0$. Notice that $\mathfrak{b}(f_+, f_-) \ge 0$ for any function f.

We have completed our main task in using the estimate (8.2) to control local singularity of B_{ij}. The rest of the proof for Theorem 7.7 is essentially the same as in [10]. For the convenience of the reader, we present the detailed proof. Besides providing some details and tracking the dependence of various constants and exponents for our application, we follow the equations in [10] very closely.

8.2.2 Time slices when the good term helps

Let Σ be the time that $u(T)$ is substantially below $\varphi^{(0)}$, i.e.,

$$\Sigma := \Big\{ T \in (-3M, -2M) \ : \ \#\Big\{|i| \le M \ : \ u_i(T) \le \varphi_i^{(0)}\Big\} \ge \frac{1}{4}\mu M \Big\}.$$

We have from (8.42) that

$$|\Sigma| \geq \frac{1}{4} M\mu \tag{8.49}$$

By (8.48), we have

$$\begin{aligned} C\lambda^2\ell^2 M &\geq \int_\Sigma \mathfrak{b}[(u(t)-\varphi^{(1)})_+, (u(t)-\varphi^{(1)})_-]\mathrm{d}t \\ &\geq -\int_\Sigma \sum_{ij}(u_i(t)-\varphi_i^{(1)})_+ B_{ij}(t)(u_j(t)-\varphi_j^{(1)})_-\mathrm{d}t \\ &\geq -cM^{-2}\int_\Sigma \sum_{ij}(u_i(t)-\varphi_i^{(1)})_+(u_j(t)-\varphi_j^{(1)})_-\mathrm{d}t \end{aligned} \tag{8.50}$$

where we have used that $u_i(t)-\varphi_i^{(1)}$ is supported on $|i| \leq 3M$ and

$$B_{ij}(t) \geq \bar{c}M^{-2}, \qquad |i|,|j| \leq 3M, \tag{8.51}$$

with some positive constant $\bar{c}$ (this follows from (8.4) and $M \gg K^\xi$). For $t \in \Sigma$ the number of j's in $|j| \leq M$ such that $u_j(t) \leq \varphi_j^{(0)}$ is at least $\frac{1}{4}\mu M$; for those j's we have

$$-(u_j(t)-\varphi_j^{(1)})_- \geq \varphi_j^{(1)} - \varphi_j^{(0)} \geq 1-\lambda^2 \geq \frac{1}{2}.$$

Thus we can bound (8.50) by

$$\geq cM^{-1}\frac{\mu}{8}\int_\Sigma \sum_i (u_i(t)-\varphi_i^{(1)})_+\mathrm{d}t \geq cM^{-1}\frac{\mu}{8\lambda}\int_\Sigma \sum_i (u_i(t)-\varphi_i^{(1)})_+^2\mathrm{d}t$$

where we have used that $(u_i(t)-\varphi_i^{(1)})_+ \leq \lambda$.

Altogether we have proved

$$\int_\Sigma \sum_i (u_i(t)-\varphi_i^{(1)})_+^2\mathrm{d}t \leq C\lambda^3\mu^{-1}\ell^2M^2 \leq \lambda^{3-\frac{1}{8}}\ell^2M^2$$

if λ is sufficiently small (recall that μ is fixed). Thus there exists a subset $\Theta \subset \Sigma$ such that

$$|\Theta| \leq \lambda^{1/8}M$$

and we have

$$\sum_i (u_i(t)-\varphi_i^{(1)})_+^2 \leq \lambda^{3-\frac{1}{4}}\ell^2 M, \qquad \forall t \in \Sigma \setminus \Theta.$$

Choosing λ small and recalling (8.49) we see that

$$\sum_i (u_i(t)-\varphi_i^{(1)})_+^2 \leq \lambda^{3-\frac{1}{4}}\ell^2 M \tag{8.52}$$

holds on a set of times t's in $[-3M,-2M]$ of measure at least $M\mu/8$.

8.2.3 Finding the intermediate set

Suppose that (8.43) is not satisfied, then there is a $T_0 \in (-2M, 0)$ such that

$$\#\Big\{i \ : \ (u_i(T_0) - \varphi_i^{(2)})_+ > 0\Big\} \geq \frac{1}{2}M\delta \tag{8.53}$$

and choose a $T_1 \in \Sigma$ (then $T_1 < T_0$) such that

$$H(T_1) = \sum_i (u_i(T_1) - \varphi_i^{(1)})_+^2 \leq \lambda^{3-\frac{1}{4}}\ell^2 M \tag{8.54}$$

(such T_1 exists by the conclusion of the previous section, (8.52)).

We also have

$$\begin{aligned} H(T_0) = \sum_i (u_i(T_0) - \varphi_i^{(1)})_+^2 &\geq \sum_i (\varphi_i^{(2)} - \varphi_i^{(1)})_+^2 \mathbf{1}\big((u_i(T_0) - \varphi_i^{(2)})_+ > 0\big) \\ &\geq \sum_i \ell^2(\lambda - \lambda^2)^2 F_i^2 \mathbf{1}\big((u_i(T_0) - \varphi_i^{(2)})_+ > 0\big) \geq C_F \frac{\lambda^2}{4}\ell^2\delta^3 M \end{aligned} \tag{8.55}$$

with some positive constant C_F. This follows from (8.53); notice first that the set (8.53) must lie in $[-3M, 3M]$ (see (8.40) and (8.41)), and even if the whole set (8.53) is near the "corner" (i.e. close to $i \sim \pm 3M$), still the sum of these F_i's is of order $\delta^3 M$ since F_i is linear near the endpoints $i = \pm 3M$.

Choose now λ small enough (depending on the fixed δ) s.t.

$$\lambda^{3-\frac{1}{4}}\ell^2 M \leq \frac{1}{16}C_F \lambda^2 \ell^2 \delta^3 M$$

Since $H(T)$ is continuous and it goes from a small value $H(T_1) \leq \frac{1}{16}C_F\lambda^2\ell^2\delta^3 M$ to a large value $H(T_0) \geq \frac{1}{4}C_F\lambda^2\ell^2\delta^3 M$, the set of intermediate times

$$D := \Big\{t \in (T_1, T_0) \ : \ \frac{1}{16}C_F\lambda^2\ell^2\delta^3 M < H(t) < \frac{1}{4}C_F\lambda^2\ell^2\delta^3 M\Big\}$$

is non-empty.

Lemma 8.6 *D contains an interval of size at least $c\delta^3 M$ with some positive constant $c > 0$. Moreover, for any $t \in D$, we have*

$$\#\big\{i \ : \ \varphi_i^{(2)} \leq u_i(t)\big\} \leq \frac{1}{2}\delta M. \tag{8.56}$$

Proof. By continuity, there is an intermediate time $T' \in D \subset [T_1, T_0]$ such that $H(T') = \frac{1}{8}C_F\lambda^2\ell^2\delta^3 M$. We can assume that T' is the largest such time, i.e.

$$H(t) > \frac{1}{8}C_F\lambda^2\ell^2\delta^3 M \qquad \text{for any} \quad t \in [T', T_0] \cap D. \tag{8.57}$$

Let $T'' = T' + c\delta^3 M$ with a small $c > 0$. We claim that $[T', T''] \subset D$. For any $t \in [T', T'']$ we can use (8.48):

$$H(t) \leq H(T') + C\ell^2\lambda^2(t - T') \leq \frac{1}{8}C_F\lambda^2\ell^2\delta^3 M + Cc\ell^2\lambda^2\delta^3 M < \frac{1}{4}C_F\lambda^2\ell^2\delta^3 M \tag{8.58}$$

if c is sufficiently small. This means that as t runs through $[T', T'']$, $H(t)$ has not reached $\frac{1}{4}C_F\lambda^2\ell^2\delta^3 M$, in particular $[T', T''] \subset (T_1, T_0)$ since $H(T_0)$ is already above this threshold. Combining then (8.58) with (8.57), we get $[T', T''] \subset D$. This proves the first statement of the lemma.

For the second statement, we argue by contradiction. Suppose we have $\#\{i \,:\, \varphi_i^{(2)} \le u_i(\tau)\} > \frac{1}{2}\delta M$ for some $\tau \in D$. Going through the estimate (8.55) for $T_0 = \tau$, we would get $H(\tau) \ge C_F\frac{\lambda^2}{4}\ell^2\delta^3 M$, but this contradicts to $\tau \in D$. This completes the proof of the lemma. $\square$

Define the exceptional set $\mathcal{F} \subset D$ of times where u is below $\varphi^{(0)}$, i.e.

$$\mathcal{F} := \Big\{ t \in D \,:\, \#\big\{|j| \le 2M \,:\, u_j(t) - \varphi_j^{(0)} \le 0\big\} \ge \mu M \Big\}$$

This set is very small, since from (8.48) we have

$$\begin{aligned}
C\lambda^2\ell^2 M &\ge -\int_{-3M}^{0}\sum_{ij}(u_i(t)-\varphi_i^{(1)})_+ B_{ij}(t)(u_j(t)-\varphi_j^{(1)})_-\mathrm{d}t \\
&\ge -\int_{\mathcal{F}}\sum_{|i|,|j|\le 3M}(u_i(t)-\varphi_i^{(1)})_+ B_{ij}(t)(u_j(t)-\varphi_j^{(1)})_-\mathrm{d}t \\
&\ge -\bar{c}M^{-2}\int_{\mathcal{F}}\sum_{|i|,|j|\le 3M}(u_i(t)-\varphi_i^{(1)})_+(u_j(t)-\varphi_j^{(1)})_-\mathrm{d}t \\
&\ge \frac{\bar{c}}{2M}\ell\mu\int_{\mathcal{F}}\sum_{|i|\le 3M}(u_i(t)-\varphi_i^{(1)})_+\mathrm{d}t
\end{aligned}$$

where we restricted the time integration to $\mathcal{F}$ in the first step, then we used (8.51) in the second step. In the third step we used that whenever $u_j(t) - \varphi_j^{(0)} \le 0$ (see the definition of $\mathcal{F}$), then $-(u_j(t) - \varphi_j^{(1)})_- \ge \ell(1-\lambda) \ge \frac{\ell}{2}$.

By (8.41), $(u_i(t) - \varphi_i^{(1)})_+ \le \ell\lambda$ and $(u_i(t) - \varphi_i^{(1)})_+ = 0$ if $|i| \ge 3M$. Hence we can continue the above estimate

$$C\lambda^2\ell^2 M \ge \frac{\bar{c}\mu}{2M\lambda}\int_{\mathcal{F}}\sum_i (u_i(t)-\varphi_i^{(1)})_+^2\mathrm{d}t = \frac{\bar{c}\mu}{2M\lambda}\int_{\mathcal{F}}H(t)\mathrm{d}t \ge \frac{\bar{c}C_F}{32}\lambda\ell^2\delta^3\mu|\mathcal{F}|$$

Here we used that $\mathcal{F} \subset D$ and in D we have a lower bound on $H(t)$. The conclusion is that

$$|\mathcal{F}| \le \frac{C\lambda}{\delta^3\mu}M$$

with some fixed constant C. Using that $|D| \ge c\delta^3 M$ from Lemma 8.6, we thus have

$$|\mathcal{F}| \le \frac{|D|}{2}$$

if λ is sufficiently small (like $\lambda \le c\delta^6\mu$).

This means that $|D \setminus \mathcal{F}| \ge \frac{c}{2}\delta^3 M$. Now we claim that for $t \in D \setminus \mathcal{F}$ we have

$$A(t) := \#\Big\{ i \,:\, \varphi_i^{(0)} < u_i(t) < \varphi_i^{(2)} \Big\} \ge \frac{M}{2}. \tag{8.59}$$

This is because $t \notin \mathcal{F}$ guarantees that the lower bound $\varphi_i^{(0)} \le u_i(t)$ is violated not more than $\mu M \le M/4$ times among the indices $|i| \le 2M$. By (8.56), the upper bound $u_i(t) \le \varphi_i^{(2)}$ is violated not more than $\frac{1}{2}\delta M \le M/4$ times.

Finally, integrating (8.59) gives

$$\int_{-3M}^{0} \#\Big\{i \ : \ \varphi_i^{(0)} < u_i(t) < \varphi_i^{(2)}\Big\}\mathrm{d}t = \int_{-3M}^{0} A(t)\mathrm{d}t \ge \frac{M}{2}|D \setminus \mathcal{F}| \ge c\delta^3 M^2.$$

with some small $c > 0$, which implies (8.44). This proves Lemma 8.5. $\square$

8.3 Oscillation estimate

For $\varepsilon > 0, \lambda > 0$ define

$$\widehat{\psi}_i = \widehat{\psi}_i^{(M,Z,\ell,\lambda,\varepsilon)} := \ell\Big[\Big(\Big|\frac{i-Z}{M}\Big| - \lambda^{-4}\Big)^{\varepsilon} - 1\Big]_+ \qquad \text{for} \qquad |i - Z| \ge M\lambda^{-4}$$

and $\widehat{\psi}_i = 0$ otherwise. This is essentially the same function as $\widetilde{\psi}$, but the growth at infinity is not $|i|^{1/4}$ but only $|i|^{\varepsilon}$. Define for any set S and any function f the oscillation $\mathrm{Osc}_S f := \sup_S f - \inf_S f$.

Lemma 8.7 *Let $Z \in [\![M, K-M]\!]$ and $\ell > 0$. Let $M^{\vartheta} = K$ and assume that* (8.2), (8.3), (8.36) *and* (8.4) *hold with some ρ and ξ so that* (8.6) *holds. Let $\lambda > 0$ be given in Lemma 8.5 and use this λ in the definition of $\widehat{\psi}$. Then there exist $\varepsilon > 0$ and $\lambda^* > 0$ such that for any solution satisfying*

$$-\ell - \widehat{\psi}_i \le u_i(t) \le \ell + \widehat{\psi}_i, \quad \forall i,\ t \in [-3M, 0], \qquad \textit{with } \widehat{\psi} = \widehat{\psi}^{(M,Z,\ell,\lambda,\varepsilon)} \tag{8.60}$$

we have

$$Osc_{Q_M} u \le (2 - \lambda^*)\ell, \qquad Q_M := [-M, 0] \times [-M, M]. \tag{8.61}$$

The constants ε and λ^* depends only on μ, γ, λ from Lemma 8.5, δ from Corollary 8.3 and ε_0 from Lemma 8.1. All these constants eventually depend only on ε_0 provided that $K \ge K_0(\vartheta, \Phi)$ with Φ satisfies (8.6).

Proof. This proof closely follows the original proof of [10]. We can set $Z = 0$. Without loss of generality, we can assume

$$\frac{1}{M^2}\int_{-3M}^{-2M} \#\Big\{i \ : \ |i| \le M,\ u_i(t) < \varphi_i^{(0)}\Big\}\mathrm{d}t \ge \mu \tag{8.62}$$

(Otherwise we can take $-u$). Let $x = [|i/M| - \lambda^{-4}]^{\varepsilon}$ and $a = 1/(4\varepsilon) \gg 1$. Then for $x \ge 1$,

$$\widehat{\psi}_i = \ell(x-1) \le \ell\frac{(x^a - 1)}{a} = 4\varepsilon\widetilde{\psi}_i \tag{8.63}$$

Choose k_0 such that $\lambda^{2k_0} = \varepsilon$. Then we have

$$\lambda^{-2k_0}\widehat{\psi}_i \le \widetilde{\psi}_i \tag{8.64}$$

For ε sufficiently small so that $k_0 = 100\gamma^{-1}$ where γ is from Lemma 8.5.

For $k \le k_0$, define $u^{(k+1)}$ by the following equation:

$$u^{(k+1)} - \ell = \frac{1}{\lambda^2}(u^{(k)} - \ell), \qquad u^{(0)} = u. \tag{8.65}$$

By (8.64), the upper bound in (8.60) becomes

$$u_i^{(k)}(t) \le \ell + \frac{1}{\lambda^{2k}}\widehat{\psi}_i \le \ell + \tilde{\psi}_i \qquad t \in [-3M, 0],\ \forall\, i, \quad k \le k_0. \tag{8.66}$$

Notice that for any $|i| \le M$ the sequence $u_i^{(k)}$ in k is decreasing, in particular $u_i^{(k)} \le \ell$. This follows from induction and from (8.60). Moreover, from (8.62) we have

$$\frac{1}{M^2}\int_{-3M}^{-2M} \#\Big\{i \ :\ |i| \le M,\ u_i^{(k)}(t) < \varphi_i^{(0)}\Big\} \mathrm{d}t \ge \mu \tag{8.67}$$

since the set in question is increasing in k.

We wish to apply Lemma 8.5 to $u^{(k)}$ which is an affine transformation of a solution u. As we note in Remark 8.4, subtracting a constant from a solution no longer yields a solution to the equation (7.57). However, since W_i is very small away from the edges by assumption (8.36), all estimates carry through with a negligible errors. Hence we can apply Lemma 8.5 to $u^{(k)}$ to get that if

$$\frac{1}{M^2}\int_{-2M}^{0} \#\Big\{i \ :\ \ u_i^{(k)}(t) > \varphi_i^{(2)}\Big\} \mathrm{d}t > \delta \tag{8.68}$$

then

$$\frac{1}{M^2}\int_{-3M}^{0} \#\Big\{i \ :\ \ \varphi_i^{(0)} < u_i^{(k)}(t) < \varphi_i^{(2)}\Big\} \mathrm{d}t \ge \gamma. \tag{8.69}$$

Therefore

$$\frac{1}{M^2}\int_{-3M}^{0} \#\Big\{i \ :\ \ u_i^{(k)}(t) > \varphi_i^{(2)}\Big\} \mathrm{d}t \le \frac{1}{M^2}\int_{-3M}^{0} \#\Big\{i \ :\ \ u_i^{(k)}(t) > \varphi_i^{(0)}\Big\} \mathrm{d}t - \gamma \tag{8.70}$$

Notice that by (8.66) and $F_i = 0$ if $|i| \ge 3M$, the inequality $u_i^{(k)}(t) > \varphi_i^{(0)}$ can hold only if $|i| \le 3M$ for any $k \le k_0$. Assuming $|i| \le 3M$ and $u_i^{(k)}(t) > \varphi_i^{(0)}$, we have

$$\frac{1}{\lambda^2}(u_i^{(k-1)} - \ell) + \ell = u_i^{(k)}(t) > \varphi_i^{(0)} \tag{8.71}$$

Since $|i| \le 3M \le \lambda^{-4}M$, we have, together with (8.40), that

$$u_i^{(k-1)} \ge \lambda^2(\widetilde{\psi}_i + F_i) + \ell \ge \varphi_i^{(2)}. \tag{8.72}$$

Therefore, we can bound the last integral in (8.70) by

$$\begin{aligned} &\frac{1}{M^2}\int_{-3M}^{0} \#\Big\{i \ :\ \ u_i^{(k)}(t) > \varphi_i^{(0)}\Big\} \mathrm{d}t \\ &\le \frac{1}{M^2}\int_{-3M}^{0} \#\Big\{i \ :\ |i| \le 3M,\ u_i^{(k-1)}(t) > \varphi_i^{(2)}\Big\} \mathrm{d}t. \end{aligned} \tag{8.73}$$

We have thus proved that

$$\begin{aligned}&\frac{1}{M^2}\int_{-3M}^{0}\#\Big\{i\ :\ \ u_i^{(k)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t\\&\le\frac{1}{M^2}\int_{-3M}^{0}\#\Big\{i\ :\ |i|\le 3M,\ u_i^{(k-1)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t-\gamma.\end{aligned}\tag{8.74}$$

Iterating this estimate k times, we would get

$$\frac{1}{M^2}\int_{-3M}^{0}\#\Big\{i\ :\ \ u_i^{(k)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t\le\frac{1}{M^2}\int_{-3M}^{0}\#\Big\{i\ :\ |i|\le 3M,\ u_i^{(0)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t-k\gamma$$

which becomes negative if, say, $k\gamma\ge 20$.

Thus there is a $k\le k_0$ such that (8.68) is violated, i.e.,

$$\frac{1}{M^2}\int_{-2M}^{0}\#\Big\{i\ :\ \ u_i^{(k)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t\le\delta.\tag{8.75}$$

Our goal is to apply Corollary 8.3 to $u^{(k+1)}$. We first check the assumptions of this Corollary. From (8.66) and the definitions of $\bar\psi$ and $\widetilde\psi$ in (8.30, 8.39), we have

$$u_i^{(k+1)}(t)\le\ell+\tilde\psi_i\le\ell+\bar\psi_i\qquad t\in[-3M,0],\ \forall\, i,\quad k+1\le k_0.\tag{8.76}$$

This verifies the first assumption (8.31) of Corollary 8.3. Since $\varphi_i^{(0)}=0$ for $|i|\le 2M$, we have

$$\begin{aligned}&\frac{1}{M^2}\int_{-2M}^{0}\#\Big\{i\ :\ |i|\le 2M,\ u_i^{(k+1)}(t)>0\Big\}\mathrm{d}t\\&=\frac{1}{M^2}\int_{-2M}^{0}\#\Big\{i\ :\ |i|\le 2M,\ u_i^{(k+1)}(t)>\varphi_i^{(0)}\Big\}\mathrm{d}t\\&\le\frac{1}{M^2}\int_{-2M}^{0}\#\Big\{i\ :\ \ u_i^{(k)}(t)>\varphi_i^{(2)}\Big\}\mathrm{d}t\le\delta,\end{aligned}$$

where we have used (8.73) in the last inequality. This verifies the second assumption (8.32) of Corollary 8.3. Thus from Corollary 8.3 we get

$$u_i^{(k+1)}(t)\le\ell/2,\qquad t\in[-M,0],\ |i|\le M.$$

Together with (8.65), we have proved

$$u_i(t)-\ell=\lambda^{2(k+1)}(u_i^{(k+1)}(t)-\ell)\le-\lambda^{2(k+1)}\ell/2,\qquad t\in[-M,0],\ |i|\le M.$$

for all $k\le k_0$. Thus $u_i(t)\le\ell(1-\lambda^*/2)$ with $\lambda^*=\frac{1}{2}\lambda^{2k_0}$ and we have proved (8.61). □

8.4 Hölder continuity

In this section we prove the Hölder regularity.

Theorem 8.8 *Let $\mathcal{M}^\vartheta = K$ with $\vartheta \geq 1$. Suppose u is a solution to* (7.57) *in the time interval* $[-3\mathcal{M}, 0]$ *with coefficients of $\mathcal{A}$ satisfy* (8.2), (8.3), (8.4) *and* (8.36) *for some $Z \in [\![L - K/2, L + K/2]\!]$ with L and K satisfying* (3.1). *Assume that u satisfies*

$$\sup_{t\in[-3\mathcal{M},0]} \max_i |u_i(t)| \leq \ell \tag{8.77}$$

for some ℓ. Then there exist constants $c_2 > 0$ and $\alpha_0 > 0$ (independent of the previous exponents ρ, ξ and ϑ) such that for any $0 < \alpha \leq \alpha_0$ we have

$$Osc_{Q^{(\alpha)}(Z)}(u) \leq \ell \mathcal{M}^{-c_2\alpha}, \qquad Q^{(\alpha)}(Z) := [-\mathcal{M}^{1-\alpha}, 0] \times [\![Z - \mathcal{M}^{1-\alpha}, Z + \mathcal{M}^{1-\alpha}]\!], \tag{8.78}$$

i.e. the oscillation of the solution on scale $\mathcal{M}^{1-\alpha}$ and away from the edges of the configuration space is smaller than $\ell\mathcal{M}^{-c_2\alpha}$. The constants c_2 and α_0 depend only on λ^* from Lemma 8.7 which depends eventually only on ε_0 from Lemma 8.1.

Proof. Given λ from Lemma 8.5, λ^* and ε from Lemma 8.7, there is a small scaling factor $\nu < (1/8)^{1/\varepsilon}$ such that

$$\left[(|\nu x| - \lambda^{-4})_+^\varepsilon - 1\right]_+ \leq \frac{1}{2}(1 - \lambda^*/2)\left[(|x| - \lambda^{-4})_+^\varepsilon - 1\right]_+$$

for any $|x| \geq \nu^{-1}$. This means that

$$\widehat{\psi}_i^{M/\nu,Z,\ell,\lambda,\varepsilon} \leq \frac{1}{2}(1 - \lambda^*/2)\widehat{\psi}_i^{M,Z,\ell,\lambda,\varepsilon} \qquad \text{if} \quad |i - Z| \geq M/\nu. \tag{8.79}$$

Set $M_k = \nu^k \mathcal{M}$, and $\ell_k = (1 - \lambda^*/4)^k \ell$, then we have

$$\frac{1}{1 - \lambda^*/4}\widehat{\psi}_i^{M_k,Z,\ell_k,\lambda,\varepsilon} \leq \frac{1}{2}\widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon} \qquad \text{if} \quad |i - Z| \geq M_k. \tag{8.80}$$

Now we define the following rescaled sequence of solutions:

$$u_i^{(0)}(t) = u_i(t), \qquad t \in [-3\mathcal{M}, 0]$$

and

$$u_i^{(k+1)}(t) = \Big(u_i^{(k)}(t) - \bar{u}^{(k)}\Big), \qquad t \in [-3M_{k+1}, 0]$$

where $\bar{u}^{(k)}$ is determined by

$$\sup_{Q_k} |u^{(k)} - \bar{u}^{(k)}| = \frac{1}{2}\mathrm{Osc}_{Q_k} u^{(k)} \tag{8.81}$$

(with $Q_k = Q_{M_k}$ for brevity).

We remind the reader that subtracting a constant from a solution no longer yields a solution. But the conclusions of previous subsections still hold as long as we made some very minor modifications to the arguments. Since this was explained in the proof of Lemma 8.7, we will not repeat it here. So from now on, we will apply these Lemmas to $u^{k)}$.

We claim that

$$\max_i |u_i^{(k)}(t)| \leq \ell_k + \widehat{\psi}_i^{M_k,Z,\ell_k,\lambda,\varepsilon}, \qquad t \in [-3M_k, 0] \tag{8.82}$$

and

$$\mathrm{Osc}_{Q_k} u^{(k)} \leq (2 - \lambda^*)\ell_k \tag{8.83}$$

for any k. By Lemma 8.7, the first statement (8.82) implies the second (8.83).

For $k = 0$, (8.82) (hence (8.83)) certainly holds by (8.77). We proceed by induction, suppose (8.82) and (8.83) hold for some k, we show (8.82) for $k+1$, i.e. we need to show that

$$|u_i^{(k+1)}(t)| = |u_i^{(k)}(t) - \bar{u}^{(k)}| \le \ell_{k+1} + \widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon}, \qquad t \in [-3M_{k+1}, 0] \tag{8.84}$$

Using (8.83), we have

$$\mathrm{Osc}_{Q_k} u^{(k)} \le (2 - \lambda^*)\ell_k \le 2\ell_{k+1} \tag{8.85}$$

so for $|i - Z| \le M_k$ we have that (8.84) is satisfied by using (8.81). Moreover, we have $|\bar{u}^{(k)}| \le \ell_k$.

Now for $|i - Z| \ge M_k$, (8.80) can be used. Moreover, in this regime we have

$$\widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon} \ge \ell_{k+1}\Big[\Big(\frac{M_k}{M_{k+1}} - \lambda^{-4}\Big)_+^{\varepsilon} - 1\Big]_+ \ge \frac{1}{2}\nu^{-\varepsilon}\ell_{k+1}$$

Choosing ν sufficiently small ($\nu^{\varepsilon} \le 1/8$) we can estimate this from below by $4\ell_{k+1}$. Thus we estimate

$$\begin{aligned} |u_i^{(k)}(t) - \bar{u}^{(k)}| \le& |u_i^{(k)}(t)| + \ell_k \le 2\ell_k + \widehat{\psi}_i^{M_k,Z,\ell_k,\lambda,\varepsilon} \\ \le& 2\ell_k + \frac{1 - \lambda^*/4}{2}\widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon}, \qquad t \in [-3M_{k+1}, 0], \quad |i - Z| \ge M_k. \end{aligned}$$

By definition, $2\ell_k \le 3\ell_{k+1}$ if λ^* is small enough. Also, we have $4\ell_{k+1} \le \widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon}$ if $|i - Z| \ge M_k$. Thus we can bound the last equation from above by

$$\le \ell_{k+1} + 2\ell_{k+1} + \frac{1}{2}\widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon} \le \ell_{k+1} + \widehat{\psi}_i^{M_{k+1},Z,\ell_{k+1},\lambda,\varepsilon}.$$

Thus we have proved (8.82) and (8.83) by induction, i.e.

$$\mathrm{Osc}_{Q_k} u^{(k)} \le \ell(1 - \lambda^*/4)^k$$

Applying it to a box of size $M_k \sim \mathcal{M}^{1-\alpha}$, we have $\nu^k = \mathcal{M}^{-\alpha}$ and thus we obtain (8.78) with

$$c_2 = \frac{\log(1 - \lambda^*/4)}{\log \nu}$$

This completes the proof. $\square$

8.5 Proof of Theorem 7.7.

With the Hölder regularity result, Theorem 8.8, we now complete the proof of Theorem 7.7. Here we return to the notations of Section 7.

From the decay estimate (7.42) with $b = K^{-\xi'}$ (see (7.60)) and $\|\mathbf{u}(0)\|_1 = 1$, we have

$$\|\mathbf{u}(s)\|_\infty \le s^{-1}K^{\xi'}. \tag{8.86}$$

As in the conditions of Theorem 7.7, let $\sigma \in [K^{c_3}, C_1 K \log K]$ be fixed and recall the relation $\mathcal{M}^{\vartheta} = K$. We now choose $\vartheta \lesssim 1/c_3$ in Theorem 8.8 so that $CM \ge \sigma \ge 9\mathcal{M}$ (recall $\mathcal{M} = 2M + 1$). Before we can apply Theorem 8.8, we have to shift the origin of the time coordinate axis by σ so

that the time interval $[-3\mathcal{M}, 0]$ in Theorem 8.8 corresponds to $[\sigma - 3\mathcal{M}, \sigma]$ in Theorem 7.7. After this time shift, (8.86) implies (8.77) if we define

$$\ell = \mathcal{M}^{-1} K^{\xi'}. \tag{8.87}$$

From the Hölder estimate, Theorem 8.8, we have

$$\mathrm{Osc}_{Q^{(\alpha)}(Z)}(u) \le K^{\xi'} \mathcal{M}^{-1-c_2\alpha} \le C\sigma^{-1+(\vartheta\xi' - c_2\alpha)}. \tag{8.88}$$

Now we choose $\alpha := \min\{\alpha_0, \frac{1}{2}q'\}$ and $q := \frac{1}{2}c_2\alpha$, where α_0 was obtained in Theorem 8.8 and q' was given in Theorem 7.7. In particular, $|j - Z| \le \sigma^{1-q'}$ implies $|j - Z| \le \mathcal{M}^{1-\alpha}$. Then, choosing ξ_0 small enough so that $\vartheta\xi_0 < q = c_2\alpha - q$, we have, since $\xi' \le \xi_0$, that

$$\mathrm{Osc}_{Q^{(\alpha)}(Z)}(u) \le \sigma^{-1-q}. \tag{8.89}$$

This proves Theorem 7.7. □

A Proof of Lemma 3.5

First we show that on the set $\mathcal{R}_{L,K}$, the length of the interval $J = J_{\mathbf{y}} = (y_{L-K-1}, y_{L+K+1})$ satisfies (3.22). We first write

$$|J| = |y_{L+K+1} - y_{L-K-1}| = |\gamma_{L+K+1} - \gamma_{L-K-1}| + O(N^{-1+\xi\delta/2}). \tag{A.1}$$

Then we use the Taylor expansion

$$\varrho(x) = \varrho(\bar{y}) + O(x - \bar{y})$$

around the midpoint $\bar{y}$ of J. Here we used that $\varrho \in C^1$ away from the edge. Thus from (2.14)

$$\mathcal{K} + 1 = N \int_{\gamma_{L-K-1}}^{\gamma_{L+K+1}} \varrho = N \int_{y_{L-K-1}}^{y_{L+K+1}} \varrho + O(N^{\xi\delta/2}) = N|J|\varrho(\bar{y}) + O(N|J|^2) + O(N^{\xi\delta/2}), \tag{A.2}$$

since the contribution of the second order term in the Taylor expansion is of order $N|J|^2$. Expressing $|J|$ from this equation and using (3.1), we arrive at (3.22).

Now we prove (3.23). We set

$$U(x) := V(x) - \frac{2}{N} \sum_{j \,:\, |j-L| \ge K+K^{\xi}} \log|x - \gamma_j|.$$

The potential U is similar to $V_{\mathbf{y}}$, but the interactions with the external points near the edges of J (y_j's with $|j - L| < K + K^{\xi}$) have been removed and the external points y_j away from the edges have been replaced by their classical value γ_j. In proving (3.23), we will first compare $V_{\mathbf{y}}$ with an auxiliary potential U and then we compute U'.

First we estimate the difference $V'_{\mathbf{y}}(x) - U'(x)$. We fix $x \in J$, and for definiteness, we assume that $d(x) = x - y_{L-K-1}$, i.e. x is closer to the lower endpoint of J; the other case is analogous.

We get (explanations will be given after the equation)

$$
\begin{aligned}
|V_{\mathbf{y}}'(x) - U'(x)| \leq & \frac{1}{N} \sum_{K<|j-L|<K+K^\xi} \frac{1}{|x-y_j|} + \frac{1}{N} \sum_{|j-L|\geq K+K^\xi} \frac{|y_j-\gamma_j|}{|x-y_j||x-\gamma_j|} \\
\leq & \frac{CK^\xi}{Nd(x)} + \frac{N^{-1+\delta\xi/2}}{d(x)} \frac{1}{N} \Big[\sum_{j=\alpha N/2}^{L-K-K^\xi} + \sum_{j=L+K+K^\xi}^{N(1-\alpha/2)} \Big] \frac{1}{|x-\gamma_j|} \\
& + \frac{CN^{-4/15+\varepsilon}}{N} \Big[\sum_{j=N^{3/5+\varepsilon}}^{\alpha N/2} 1 + \sum_{N(1-\alpha/2)}^{N-N^{3/5+\varepsilon}} 1 \Big] + \frac{C}{N} \Big[\sum_{j=1}^{N^{3/5+\varepsilon}} 1 + \sum_{j=N-N^{3/5+\varepsilon}}^{N} 1 \Big] \\
\leq & \frac{CK^\xi}{Nd(x)} + \frac{CN^{-1+\delta\xi/2}\log N}{d(x)} + CN^{-4/15+\varepsilon} \\
\leq & \frac{CK^\xi}{Nd(x)}.
\end{aligned} \tag{A.3}
$$

Here for the first bulk sum, $j \in [\![N\alpha/2, L-K-K^\xi]\!]$, we used $|y_j - \gamma_j| \leq N^{-1+\xi\delta/2}$ from the definition of $\mathcal{R}_{L,K}$ and the fact that for $j \leq L-K-K^\xi$ we have

$$
\begin{aligned}
x - \gamma_j & \geq y_{L-K-1} - \gamma_j \\
& \geq \gamma_{L-K-1} - \gamma_j - |y_{L-K-1} - \gamma_{L-K-1}| \\
& \geq cN^{-1}(L-K-1-j) - CN^{-1+\xi\delta/2} \\
& \geq c'N^{-1}(L-K-1-j)
\end{aligned}
$$

with some positive constants c, c'. This estimate allows one to sum up $|x-\gamma_j|^{-1}$ at the expense of a $\log N$ factor. Similar estimate holds for $j \geq L+K+K^\xi$. In the intermediate sum, $j \in [\![N^{3/5+\varepsilon}, N\alpha/2]\!]$, we used $|y_j - \gamma_j| \leq CN^{-4/15+\varepsilon}$ and that $|x-y_j|$ and $|x-\gamma_j|$ are bounded from below by a positive constant since

$$
x - y_j \geq y_{L-K-1} - y_j \geq y_{\alpha N} - y_j \geq \gamma_{N\alpha} - \gamma_{N\alpha/2} + O(N^{-1+\xi\delta/2}) \geq c
$$

and similarly for $x-\gamma_j$. Finally, very near the edge, e.g. for $j \leq N^{3/5+\varepsilon}$, we just estimated $|y_j - \gamma_j|$ by a constant. This explains (A.3).

Now we estimate $U'(x)$. We use the fact that the equilibrium measure $\varrho = \varrho_V$ satisfies the identity

$$
\frac{1}{2}V'(x) = \int \frac{\varrho(y)}{x-y} \mathrm{d}y
$$

from the Euler-Lagrange equation of (2.13), see [1, 9]. Thus

$$
\frac{1}{2}|U'(x)| \leq |\Omega_1| + |\Omega_2| + |\Omega_3|
$$

with

$$\Omega_1 := \int_{\gamma_{L-K-K^\xi}}^{\gamma_{L+K+K^\xi}} \frac{\varrho(y)}{x-y} \mathrm{d}y,$$

$$\Omega_2 := \int_A^{\gamma_{L-K-K^\xi}} \frac{\varrho(y)}{x-y} \mathrm{d}y - \frac{1}{N} \sum_{j=1}^{L-K-K^\xi} \frac{1}{x-\gamma_j},$$

$$\Omega_3 := \int_{\gamma_{L+K+K^\xi}}^{B} \frac{\varrho(y)}{x-y} \mathrm{d}y - \frac{1}{N} \sum_{j=L+K+K^\xi}^{N} \frac{1}{x-\gamma_j},$$

where $[A, B]$ is the support of the density ρ.

To estimate Ω_1, we use Taylor expansion

$$\varrho(y) = \varrho(x) + O\big(|x-y|\big).$$

For definiteness we again assume that $d(x) = x - y_{L-K-1}$, and use that on $\mathcal{R}_{L,K}$ we have

$$\gamma_{L-K-K^\xi} \leq y_{L-K-1} \leq x \leq y_{L+K+1} \leq \gamma_{L+K+K^\xi}.$$

We thus obtain

$$\begin{aligned}
\Omega_1 = & \int_{\gamma_{L-K-K^\xi}}^{\gamma_{L+K+K^\xi}} \frac{\varrho(x) + O\big(|x-y|\big)}{x-y} \mathrm{d}y \\
= & \varrho(x) \log \frac{\gamma_{L+K+K^\xi} - x}{x - \gamma_{L-K-K^\xi}} + O(K/N) \\
= & \varrho(\bar{y}) \log \frac{d_+(x)}{d_-(x)} + O(KN^{-1+\xi}).
\end{aligned} \tag{A.4}$$

In the first step above we computed the leading term of the integral, while the other term was estimated trivially using that the integration length is $\gamma_{L+K+K^\xi} - \gamma_{L-K-K^\xi} = O(K/N)$. In the second step we used that $\varrho \in C^1$ away the edge, i.e. $\varrho(x) = \varrho(\bar{y}) + O(K/N)$. To estimate the logarithm, we used

$$\begin{aligned}
\gamma_{L+K+K^\xi} - x = & (\gamma_{L+K+K^\xi} - \gamma_{L+K+1}) + (\gamma_{L+K+1} - y_{L+K+1}) + (y_{L+K+1} - x) \\
= & \varrho(\bar{y}) N^{-1} K^\xi + O(N^{-1+\xi\delta/2}) + (y_{L+K+1} - x) \\
= & d_+(x) + O(N^{-1+\xi\delta/2})
\end{aligned}$$

and the similar relation

$$x - \gamma_{L-K-K^\xi} = d_-(x) + O(N^{-1+\xi\delta/2}).$$

Notice that the error term in (A.4) is smaller than the target estimate $K^\xi/(Nd(x))$ since $d(x) \leq K/N \ll K^{-1+\xi} N^{-\xi}$.

Now we estimate the Ω_2 term; Ω_3 can be treated analogously. We can write (with the convention

$\gamma_0 = A$)

$$\begin{aligned}|\Omega_2| =& \Big|\sum_{j=1}^{L-K-K^\xi} \int_{\gamma_{j-1}}^{\gamma_j} \varrho(y)\Big[\frac{1}{x-y} - \frac{1}{x-\gamma_j}\Big]\mathrm{d}y\Big| \\ \leq& C\sum_{j=1}^{L-K-K^\xi}(\gamma_j - \gamma_{j-1})\int_{\gamma_{j-1}}^{\gamma_j}\frac{\varrho(y)}{(x-y)^2}\mathrm{d}y \\ \leq& CN^{-1}\int_{A+\kappa}^{\gamma_{L-K-K^\xi}}\frac{\mathrm{d}y}{(x-y)^2} + CN^{-2/3}\int_A^{A+\kappa}\frac{\mathrm{d}y}{(x-y)^2} \\ \leq& \frac{CN^{-1}}{d(x)}.\end{aligned}$$

In the first step we used that

$$\int_{\gamma_{j-1}}^{\gamma_j} \varrho(y) = \frac{1}{N}$$

from (2.14). In the second step we used that $\gamma_j - \gamma_{j-1} = O_\kappa(N^{-1})$ in the bulk, i.e. for $\gamma_j \geq A+\kappa$, and $\max_j(\gamma_j - \gamma_{j-1}) = O(N^{-2/3})$ (the order $N^{-2/3}$ comes from the fact that the density ρ vanishes as a square root at the endpoints). The parameter $\kappa = \kappa(\alpha)$ is chosen such that $A + 2\kappa \leq y_{L-K-1}$ which can be achieved since $L \geq \alpha N$ and y_{L-K-1} is close to γ_{L-K-1}. In the very last step we absorbed the $N^{-2/3}$ error term into $(Nd(x))^{-1} \geq K^{-1} \gg N^{-2/3}$.

Finally we prove (3.24). Since $|y_j - \gamma_j| \leq K^\xi/N$, it follows that $|x-y_j| \sim |x-\gamma_j|$ for $|x-\gamma_j| \geq K^\xi/N$. Thus we have

$$V_{\mathbf{y}}''(x) = V''(x) + \frac{2}{N}\sum_{j\notin I}\frac{1}{(x-y_j)^2} \geq \inf V'' + \frac{c}{N}\sum_{j\notin I}\frac{1}{(x-\gamma_j)^2} \geq \inf V'' + \frac{c}{d(x)},$$

with some positive constant c (depending only on α). In estimating the summation, we used that the sequence γ_k is regularly spaced with gaps of order $1/N$. This completes the proof of Lemma 3.5. $\square$

B Critical discrete Gagliardo-Nirenberg inequality for $|p|$

Proposition B.1 *There exists a positive constant C such that*

$$\|f\|_{L^4(\mathbb{Z})}^4 \leq C\|f\|_{L^2(\mathbb{Z})}^2 \sum_{i\neq j\in\mathbb{Z}} \frac{|f_i - f_j|^2}{|i-j|^2} \tag{B.1}$$

holds for any function $f : \mathbb{Z} \to \mathbb{R}$.

Proof. Recall the integral formula for quadratic form of the operator $|p|$ in $\mathbb{R}$ [39]:

$$\int_{\mathbb{R}} \phi(x)\,(|p|\,\phi)(x)\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{|\phi(x)-\phi(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y. \tag{B.2}$$

We have the the following Gagliardo-Nirenberg type inequality in the critical case (see (1.4) of [42] with the choice $n = 1, p = 4$)

$$\|\phi\|_4^4 \le C\|\phi\|_2^2 \int_{\mathbb{R}} \phi(x)\,(|p|\,\phi)(x)\mathrm{d}x, \qquad \phi : \mathbb{R} \to \mathbb{R}. \tag{B.3}$$

Thus (B.1) is a discrete version of (B.3). Now we show how to to derive (B.1) from (B.3).

Given $f : \mathbb{Z} \to \mathbb{R}$, let $\phi : \mathbb{R} \to \mathbb{R}$ be its linear interpolation, i.e. $\phi(i) := f_i$ for $i \in \mathbb{Z}$ and

$$\phi(x) = f_i + (f_{i+1} - f_i)(x - i) = f_{i+1} - (f_{i+1} - f_i)(i + 1 - x), \qquad x \in [i, i + 1]. \tag{B.4}$$

It is easy to see that

$$C^{-1}\|\phi\|_{L^p(\mathbb{R})} \le \|f\|_{L^p(\mathbb{Z})} \le C\|\phi\|_{L^p(\mathbb{R})}\ , \qquad 2 \le p \le 4, \tag{B.5}$$

with some universal constant C. We claim that

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y \le C \sum_{i\neq j\in\mathbb{Z}} \frac{|f_i - f_j|^2}{|i - j|^2} \tag{B.6}$$

with a universal constant C, then (B.5) and (B.6) will yield (B.1) from (B.3).

To prove (B.6), we can write

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y = \sum_{i,j} \int_i^{i+1}\int_j^{j+1} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y \tag{B.7}$$

Using the explicit formula (B.4), we first compute the $i = j$ terms in (B.7):

$$\sum_i \int_i^{i+1}\int_i^{i+1} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y = \sum_i |f_i - f_{i+1}|^2 = \sum_i \frac{|f_i - f_{i+1}|^2}{|i - (i + 1)|^2}. \tag{B.8}$$

Next we compute the terms $|i - j| = 1$ in (B.7). We assume $j = i - 1$, the terms $j = i + 1$ are analogous;

$$\begin{aligned} &\sum_i \int_i^{i+1}\int_{i-1}^{i} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y \qquad\qquad\qquad\qquad (\mathrm{B.9})\\ &\quad\le \sum_i (f_{i+1} - f_i)^2 \int_i^{i+1}\int_{i-1}^{i} \frac{(x - i)^2}{(x - y)^2}\mathrm{d}x\mathrm{d}y + \sum_i (f_i - f_{i-1})^2 \int_i^{i+1}\int_{i-1}^{i} \frac{(i - y)^2}{(x - y)^2}\mathrm{d}x\mathrm{d}y, \end{aligned}$$

where we used $\phi(x) = f_i + (f_{i+1} - f_i)(x - i)$ and $\phi(y) = f_i - (f_i - f_{i-1})(i - y)$. The above integrals are finite constants, so we get

$$\sum_i \int_i^{i+1}\int_{i-1}^{i} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y \le C\sum_i \frac{(f_{i+1} - f_i)^2}{(i + 1 - i)^2} + \frac{(f_i - f_{i-1})^2}{(i - (i - 1))^2}.$$

Finally, for the terms $|i - j| \ge 2$, we can just replace $(x - y)^2$ by $(i - j)^2$ in the right hand side of (B.7) and use simple Schwarz inequalities to get

$$\int_i^{i+1}\int_j^{j+1} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\mathrm{d}x\mathrm{d}y \le C\frac{|f_i - f_j|^2 + |f_{i+1} - f_i|^2 + |f_{j+1} - f_j|^2}{|i - j|^2}.$$

After summing up we get

$$\sum_{|i-j|\geq 2}\int_i^{i+1}\int_j^{j+1}\frac{|\phi(x)-\phi(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y \leq C\sum_{|i-j|\geq 2}\frac{|f_i-f_j|^2}{|i-j|^2}+C\sum_i\frac{|f_{i+1}-f_i|^2}{((i+1)-i)^2}. \tag{B.10}$$

The estimates (B.8), (B.9) and (B.10) together yield (B.6). □

References

[1] Albeverio, S., Pastur, L., Shcherbina, M.: On the $1/n$ expansion for some unitary invariant ensembles of random matrices, *Commun. Math. Phys.* **224**, 271–305 (2001).

[2] Bach, V.: Moller, J.-S.: Correlation at low temperature. I. Exponential decay. *J. Funct. Anal.* **203** (2003), no.1. 93–148.

[3] Ben Arous, G., Bourgade, P.: Extreme gaps between eigenvalues of random matrices. Preprint. arXiv:1010.1294

[4] Ben Arous, G., Guionnet, A.: Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields* **108**, 4 (1997), 517–542.

[5] Ben Arous, G., Péché, S.: Universality of local eigenvalue statistics for some sample covariance matrices. *Comm. Pure Appl. Math.* **LVIII.** (2005), 1–42.

[6] Bleher, P., Its, A.: Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. *Ann. of Math.* **150**, 185–266 (1999).

[7] P. Bourgade, L. Erdős, H.-T. Yau: *Universality of General β-Ensembles.*
`(http://arxiv.org/abs/1104.2272)`

[8] P. Bourgade, L. Erdős, H.-T. Yau: *Bulk Universality of General β-Ensembles with Non-convex Potential.*
`(http://arxiv.org/abs/1201.2283)`

[9] Boutet de Monvel, A., Pastur, L., Shcherbina, M.: On the statistical mechanics approach in the Random Matrix Theory. Integrated Density of States. *J. Stat. Phys.* **79** , 585–611 (1995).

[10] Caffarelli, L., Chan, C.H., Vasseur, A.: Regularity theory for parabolic nonlinear integral operators, *J. Amer. Math. Soc.* **24**, No. 3 (2011), 849–889.

[11] Deift, P.: Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. *Courant Lecture Notes in Mathematics* **3**, American Mathematical Society, Providence, RI, 1999.

[12] Deift, P., Gioev, D.: Universality in random matrix theory for orthogonal and symplectic ensembles. Int. Math. Res. Pap. IMRP 2007, no. 2, Art. ID rpm004, 116 pp

[13] Deift, P., Gioev, D.: Random Matrix Theory: Invariant Ensembles and Universality. *Courant Lecture Notes in Mathematics* **18**, American Mathematical Society, Providence, RI, 2009.

[14] Deift, P., Kriecherbauer, T., McLaughlin, K.T-R, Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52**, 1335–1425 (1999).

[15] Deift, P., Kriecherbauer, T., McLaughlin, K.T-R, Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52**, 1491–1552 (1999).

[16] Deuschel, J.-D., Giacomin, G., Ioffe, D.: Large deviations and concentration properties for $\nabla\varphi$ interface models. *Probab. Th. Relat. Fields.* **117** (2000), 49–111.

[17] Dumitriu, I., Edelman, A.: Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002), 5830–5847.

[18] Dyson, F.J.: A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3**, 1191–1198 (1962).

[19] Erdős, L., Péché, S., Ramirez, J., Schlein, B., Yau, H.-T.: Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63**, No. 7, 895–925 (2010)

[20] Erdős, L., Ramírez, J., Schlein, B., Tao, T., Vu, V. and Yau, H.-T., Bulk universality for Wigner hermitian matrices with subexponential decay. Math. Res. Lett. **17** (2010), no. 4, 667–674.

[21] Erdős, L., Schlein, B., Yau, H.-T.: Universality of Random Matrices and Local Relaxation Flow. Invent. Math. **185** (2011), no.1, 75–119.

[22] Erdős, L., Schlein, B., Yau, H.-T., Yin, J.: The local relaxation flow approach to universality of the local statistics for random matrices. to appear in *Annales Inst. H. Poincaré (B), Probability and Statistics.* Preprint `arXiv:0911.3687`

[23] Erdős, L., Ramírez, J.-A., Schlein, B., Yau, H.-T.: Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation, *Electronic Journal of Probability,* **15** (2010), Paper no. 18, pages 526–603.

[24] L. Erdős, H.-T. Yau, J. Yin, *Bulk universality for generalized Wigner matrices.* Accepted to Prob. Theor. Rel. Fields `(http://arxiv.org/abs/1001.3453)`

[25] L. Erdős, H.-T. Yau, J. Yin, *Universality for generalized Wigner matrices with Bernoulli distribution.* J. of Combinatorics, **1** (2011), no. 2, 15–85

[26] Erdős, L., Yau, H.-T., Yin, J., Rigidity of Eigenvalues of Generalized Wigner Matrices. Accepted to Adv. Math. Preprint `arXiv:1007.4652.`

[27] Erdős, L., Yau, H.-T., Yin, J.: Bulk universality for generalized Wigner matrices. Preprint arXiv:1001.3453.

[28] Erdős, L., Yau, H.-T., Yin, J.: Universality for generalized Wigner matrices with Bernoulli distribution. Preprint arXiv:1003.3813.

[29] Erdős, L., Yau, H.-T., Universality of local spectral statistics of random matrices. `(http://arxiv.org/abs/1106.4986)` To appear in Bull. AMS

[30] Erdős, L., Yau, H.-T., A comment on the Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices. Preprint: arXiv:1201.5619.

[31] Fokas, A. S.; Its, A. R.; Kitaev, A. V.: The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.* **147** (1992), no. 2, 395–430.

[32] Giacomin, G., Olla, S., Spohn, H.: Equilibrium fluctuations for $\nabla\varphi$ interface model. *Ann. Probab.* **29** (2001), no.3., 1138–1172

[33] Gustavsson, J. : Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. H. Poincare Probab. Statist.* 41 (2005), no. 2, 151-178

[34] Helffer, B.: Semiclassical Analysis, Witten Laplacians and Statistical Mechanics, *Series on Partial Differential Equations and Applications.* Vol. 1. World Scientific, 2002.

[35] Helffer, B., Sjöstrand, J.: On the correlation for Kac-like models in the convex case. *J. Statis. Phys.* **74** (1994), no.1-2, 349–409.

[36] Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* **215** (2001), no.3. 683–705.

[37] Knowles, A. and Yin, J: Eigenvector Distribution of Wigner Matrices, Preprint `arXiv:1102.0057`

[38] Kriecherbauer, T., Shcherbina, M.: Fluctuations of eigenvalues of matrix models and their applications. Preprint `arXiv:1003.6121`

[39] Lieb, E. H.; Loss, M.: Analysis. Second edition. Graduate Studies in Mathematics, **14**. American Mathematical Society, Providence, RI, 2001

[40] Mehta, M.L.: Random Matrices. Academic Press, New York, 1991.

[41] Naddaf, A., Spencer, T.: On homogenization and scaling limit of some gradient perturbations of a massless free field, *Commun. Math. Phys.* **183** (1997), no.1., 55–84.

[42] Ogawa, T.; Ozawa, T.: Trudinger type inequalities and uniqueness of weak solutions for the nonlinear Schrdinger mixed problem. *J. Math. Anal. Appl.* **155** (1991), no. 2, 531540.

[43] Pastur, L., Shcherbina, M.: Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Stat. Phys.* **86**, 109–147, (1997)

[44] Pastur, L., Shcherbina M.: Bulk universality and related properties of Hermitian matrix models. *J. Stat. Phys.* **130**, no. 2., 205–250 (2008).

[45] Shcherbina, M.: Orthogonal and symplectic matrix models: universality and other properties. Preprint `arXiv:1004.2765`

[46] Tao, T. and Vu, V.: Random matrices: Universality of the local eigenvalue statistics. *Acta Math.*, **206** (2011), no. 1, 127–204.

[47] Tao, T.: The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. Preprint. arXiv:1203.1605

[48] Tracy, C., Widom, H.: Level spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159**, 151–174, 1994.

[49] C. Tracy, H. Widom, On orthogonal and symplectic matrix ensembles, *Comm. Math. Phys.* **177** (1996), no. 3, 727–754.

[50] Valkó, B.; Virág, B.: Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** (2009), no. 3, 463–508.

[51] Valkó, B.; Virág, B.: Large gaps between random eigenvalues. *Ann. Probab.* **38**, No. 3 (2010), 1263–1279.

[52] Widom H.: On the relation between orthogonal, symplectic and unitary matrix ensembles. *J. Statist. Phys.* **94** (1999), no. 3-4, 347–363.

[53] Wigner, E.: Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* **62** (1955), 548-564.